

Article

Not peer-reviewed version

Bayesian Analysis of Tuberculosis Spread Scenarios in Regions of Russian Federation

[Olga Krivorotko](#)*, [Andrei Neverov](#), [Yakov Schwartz](#), [Grigoriy Kaminskiy](#), [Nikolay Zyatkov](#), [Zhanna Laushkina](#)

Posted Date: 27 March 2026

doi: 10.20944/preprints202603.2218.v1

Keywords: mathematical modeling; Tuberculosis; multi-drug resistant forms; inverse problem; Sobol analysis; optimization; Bayesian approach; probabilistic forecasting



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Bayesian Analysis of Tuberculosis Spread Scenarios in Regions of Russian Federation

Olga Krivorotko^{1,*†}, Andrei Neverov¹, Yakov Schwartz², Grigoriy Kaminskiy³, Nikolay Zyatkov¹ and Zhanna Laushkina²

¹ Sirius University of Science and Technology, Krasnodar region, Russia

² Novosibirsk TB Research Institute of the Ministry of Health RF, Novosibirsk, Russia

³ Tula Regional Clinical Center for the Prevention and Control of AIDS and Infectious Diseases, Tula, Russia

* Correspondence: krivorotko.olya@mail.ru

† These authors contributed equally to this work.

Abstract

The paper proposes a new model of Tuberculosis (TB) dynamics taking into account multi-drug resistant forms, which takes into account the detection of infected people with and without bacterial excretion. The model is described by a system of nine nonlinear ordinary differential equations united by the law of mass action and controlled by 10 epidemiological parameters. The conditions for the stability of the system's equilibrium states are obtained, and the sensitivity-based identifiability analysis of the model is conducted using the Sobol method. Based on Bayesian optimization, the boundaries of sensitive parameters are specified and posterior distributions of the model parameters are obtained for five regions of the Russian Federation based on statistics from 2009 to 2020. It is shown the heterogeneity of epidemic situation by wide credible intervals of correlated parameters of virus contagiousness, the proportion of infected TB converting to the bacterial excretion form and the rate of detection of TB infected with bacterial excretion. Probabilistic forecasts of the expected number of TB infections to 2025 are constructed and validated to the 2021-2023 data.

Keywords: mathematical modeling; Tuberculosis; multi-drug resistant forms; inverse problem; Sobol analysis; optimization; Bayesian approach; probabilistic forecasting

1. Mathematical Modeling of the Tuberculosis Spread

Despite some success in the fight against Tuberculosis in the Russian Federation and in the world, the disease remains one of the main challenges to national healthcare systems, being the leading global cause of death from infectious diseases, one of the 10 main causes of death and the predominant cause of death in people infected with HIV. According to WHO, last year 10.6 million people fell ill with Tuberculosis for the first time, 1.3 million people died, the number of MDR cases increased by 3% [1,2]. For successful control over the spread of socially significant and epidemically dangerous infectious diseases such as Tuberculosis, predictive mathematical modeling of the epidemic process is necessary, capable of providing a reliable forecast of the probability of developing an unfavorable and favorable scenario of the epidemic situation in specific territories and early planning/implementation of the necessary interventions [3].

The difficulty of developing a model of slow infections dynamics (HIV / AIDS, Tuberculosis) is to consider the heterogeneity of the population with various climatic, environmental, epidemiological, medical and biological, socio-economic characteristics, including migration, demographic characteristics of regions, availability of medical services, the presence of high-risk groups (HIV-infected people, people with chronic diseases, addiction to psychoactive substances, etc.) and a high proportion of drug-resistant strains.

The models used for the dynamics of infectious diseases in populations have varying degrees of complexity, but three main approaches and their combinations can be distinguished [4]. The first

approach is to divide the population into non-overlapping, homogeneously mixed groups with respect to infection and to introduce rules for transition from one group to another. [5]. Typically, models of this type are systems of integral, integro-differential and difference equations (SIR models), interconnected by the mass action law. In the review [5] examples of SIR models and their features in describing the dynamics of Tuberculosis in a population are given. Such models are characterized by their parameters (coefficients and initial conditions), and describe an outbreak of an infectious disease if coefficients are constant. Accounting for medical and socio-economic interventions is possible in a generalized manner, while detailing the features of related processes is difficult.

In work [6] It is shown that the parameter of Tuberculosis infectiousness is sensitive to morbidity, therefore demographic and socio-economic indicators were introduced as a step function into the change of infectiousness parameter. This approach is demonstrated in predicting morbidity in the regions of the Russian Federation and confirms that the change of socio-economic conditions in the regions is one of the reasons for the decrease of Tuberculosis spread. Furthermore this effect is combined with the results of the work of the anti-Tuberculosis services. In work [7] for the model of Tuberculosis and HIV co-infection, fundamental scenarios of the disease spread were constructed with regulating the parameter of the effectiveness of treatment in the regions of the Russian Federation. However, including socio-economic processes in an explicit form requires modification of the model (introduction of additional equations) and data. The second approach is agent based modeling (e.g. [8]). Within this framework, each individual of the population is described by a set of parameters (age, immune status, time elapsed since infection, etc.), the rules of interaction of individuals and the influence of the environment. This approach allows us to obtain a more detailed picture of the epidemic process: age stratification of morbidity, the influence of social status and the environment on the incidence and prevalence of infection, changes in immunity as a result of the spread of infection, etc. To specify models of this type, both differential equations and algorithmic descriptions are used. Such models are computationally expensive and require a detailed description specific to the region under consideration, however, the results of modeling and forecasting are not only advisory, but also practical for medical organizations (e.g. [9]).

The third approach to describing and forecasting of infectious diseases spread relies on statistical data and deep learning models. However, their implementation requires complete data of the process under study, which is difficult to obtain for slowly progressing infections (for example, when modeling and forecasting COVID-19, the effectiveness of such models is studied in the works [10,11]).

To build scenarios for the spread of Tuberculosis in the regions of the Russian Federation, it is necessary to take into account the heterogeneity of its distribution [12], as well as the large percentage of multidrug-resistant forms of Tuberculosis (MDR-TB) in the Siberian regions of the Russian Federation among newly diagnosed patients (see Figure 1). For example, more than 40% of newly diagnosed patients were MDR-TB infected in the Novosibirsk region in 2024. While the average figure for the Siberian Federal District fluctuates around 30%.

The inclusion of the incipient and subclinical TB (ISbTB) compartment is crucial for modeling TB elimination, as it represents a major obstacle to ending TB alongside latent TB [13]. This form is characterized by a low, oscillating bacterial load, qualifying as paucibacillary disease [14], yet remains sufficient for transmission, especially in prolonged close contact [15]. Such individuals are significant concealed community sources of infection [16], eluding diagnosis due to low bacterial burdens that cause false-negative results [17] and present a "hard to determine" category [18].

ISbTB escalates the total epidemic contagiousness; its basic reproduction number, even if low, contributes to the continuity of the overall TB epidemic process. Furthermore, these forms can harbor drug resistance [19] and lead to relapse post-treatment [20]. Despite its proven epidemiological significance, many models omit this parameter. Our model aims to address this gap.

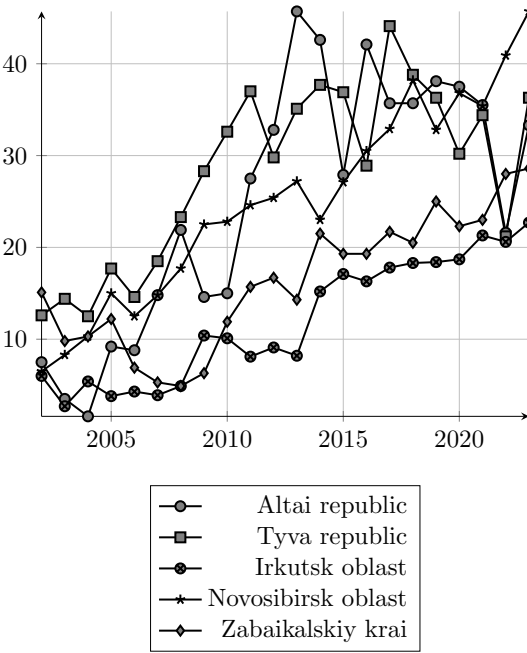


Figure 1. Proportion of MDR-TB cases among newly diagnosed patients in regions of the Russian Federation (%).

In the recent years the modeling of the epidemics is also developed with stochastic approach, also with stochastic control strategies [21,22].

The interest in mathematical modeling of TB was strongly elevated after the issue of WHO concept of ending TB by 2030 [23]. WHO tuberculosis researchers published a mathematical model to predict TB incidence decline with fulfillment of the Sustainable Development Goal (SDG) subtargets [24]. In the study [25] the susceptible-exposed-infectious-recovered (SEIR) model with different age groups was developed and the feature of heterogeneity considered in the model allowed to assess the effect of age as a factor on TB transmission and more accurately determine parameters for ending TB. Separate studies have considered the effects of drug-resistant cases [26], time lag [27] and age structure.

Questions of the hidden reservoirs of the TB infection and the burden of overall drug resistance inside them were not significantly involved in mathematical modeling of the TB infection in current literature.

It has recently become clear that traditional notions of dichotomizing tuberculosis infection into latent and active tuberculosis are a simplification that fails to reflect the biological process of transition from infection to disease. To more realistically characterize this process, the concept of two main intermediate groups of patients, existing between infection and active TB, has been introduced: incipient and subclinical TB disease. Therefore, it is now believed that the immediate precursors of the development of active TB are not so much the infection state itself (latent TB, LTBI), but rather those stages of infection progression that directly precede active TB, i.e., incipient and subclinical TB [28–31]. Recently, the first evidence has emerged of the important role of incipient and subclinical TB in the spread of infection [31,32], however, the actual contribution of these conditions to the TB epidemic remains unclear, as diagnosing them in patients is difficult. Molecular biological methods currently being developed for detecting incipient and subclinical TB will soon be introduced into healthcare practice. These methods will enable a more accurate understanding of the true contribution to the spread of TB and the prognosis of the epidemic. [33–38]. Since this contribution depends on numerous dynamic variables influencing the spread of the epidemic (human population characteristics, drug resistance of circulating strains, access to medical care, etc.), it is appropriate to describe scenarios for the impact of incipient and subclinical TB on the epidemic process using mathematical modeling. In our paper we construct the model, for the first time, examines not LTBI, but rather impact of incipient and subclinical TB in TB epidemiology.

This mathematical model based on the comprehensive SEIS-type compartmental model for TB dynamics in the Russian Federation, incorporating multi-drug resistant (MDR) forms, detection processes, and treatment pathways. In the methods we combine differential equation modeling based on M.I. Perelman and G.I. Marchuk ideas [39] as well as Y. Yi et al. model [40], Sobol sensitivity analysis, Bayesian optimization, and MCMC based forecasting, applying the framework to several high burden regions, that are of epidemiological interest in the Russian Federation as the obstacle of ending TB in Russia. The key points of the current publication, that allow to achieve additional results, as comparing with the other authors:

- Parameter identification and optimization with the use of the concept and paradigm of inverse and ill-posed problems allowed us to evaluate the intensity of epidemic process in the risk regions of Russia, that we need to identify with the aim to combat TB as the regional and national high stream.
- Sensitivity-based identifiability analysis allows us to find correlated TB parameters that implicate on the accuracy of parameter identifiability and forecast uncertainty.
- We construct the posterior distribution of sensitive epidemiological parameters of the TB compartmental mathematical model (such as contagiousness of TB contact with bacterioexcretion, the rate of TB activation, the rate of undetected TB contact with bacterioexcretion per year) that allows us to evaluate the expected TB infected people in Russian Federation regions for three years ahead.

If we get robust estimations of epidemic force parameters, we have enough scientific potential to plan counterforce actions including efficient treatment of drug resistance and drug sensitive forms of TB in the amounts enough to concur transmission. The results of the expected heterogeneity of regions and the assessment of key epidemiological parameters are consistent with the results of the works of A.A. Romanyukha [9] and O.A. Melnichenko [6], and also expand and clarify the uncertainty due to the inclusion of incipient and subclinical TB forms, i.e. the key ideas of this paper are

- Inclusion of detection stages based on papers [6,9] and add new compartments of MDR-TB based on [40];
- Inclusion of incipient/subclinical TB as a compartment based on [41];
- Usage of Sobol sensitivity analysis and Bayesian MCMC for uncertainty quantification;
- Region-specific posterior parameter estimation.

The paper is organized as follows: in the section 2 we formulate the differential SIR-type model, give analysis of stability of stationary points of the model, formulate the problem of parameters identification and provide sensitivity analysis. In the section 3 the statistical data used on Tuberculosis incidence, socio-economic characteristics of the regions of the Russian Federation and their processing for use in modeling are presented. In the section 4 numerical estimates of the probability distribution of epidemiological parameters based on statistical data and SIR model data, as well as scenarios for the spread of Tuberculosis in the regions of the Russian Federation are presented.

2. SEIS Model of Tuberculosis Dynamics with Multiple Drug-Resistant Forms

The mathematical model of Tuberculosis (TB) spread and control used in this paper is based on the approach proposed by G.I. Marchuk and A.A. Romanyukha in 2004 [39] and paper by A.A. Romanyukha [41] concerning the TB detection. Using the law of mass action, we formulate the SEIS model [40], characterized by a system of 9 ordinary differential equations (the model diagram is shown in Figure 2):

$$\left\{\begin{array}{l} \frac{dS}{dt} = -\beta S(I_s + I_\gamma)/N + \gamma_1(E_\gamma^T + I_\gamma^T) + \gamma_2(E_s^T + I_s^T) + bN - mN^2 - \mu S; \\ \frac{dE_s}{dt} = \beta SI_s/N - \omega E_s - \varphi^- E_s - \beta E_s I_\gamma/N - \mu E_s; \\ \frac{dE_s^T}{dt} = -\gamma_2 E_s^T + \gamma_2 I_s^T + \varphi^- E_s - E_s^T(\beta I_\gamma/N + t_\gamma) - \mu E_s^T; \\ \frac{dE_\gamma}{dt} = \beta SI_\gamma/N - \omega E_\gamma - \varphi^- E_\gamma + \beta E_s I_\gamma/N - \mu E_\gamma; \\ \frac{dE_\gamma^T}{dt} = \gamma_1 I_\gamma^T - \gamma_1 E_\gamma^T + \varphi^- E_\gamma + E_s^T(\beta I_\gamma/N + t_\gamma) - \mu E_\gamma^T; \\ \frac{dI_s}{dt} = \omega E_s - \varphi^+ I_s - \beta I_s I_\gamma/N - (\alpha + \mu) I_s; \\ \frac{dI_s^T}{dt} = -2\gamma_2 I_s^T + \varphi^+ I_s - I_s^T(\beta I_\gamma/N + t_\gamma) - (\alpha + \mu) I_s^T; \\ \frac{dI_\gamma}{dt} = \omega E_\gamma - \varphi^+ I_\gamma + \beta I_s I_\gamma/N - (\alpha + \mu) I_\gamma; \\ \frac{dI_\gamma^T}{dt} = -2\gamma_1 I_\gamma^T + \varphi^+ I_\gamma + I_s^T(\beta I_\gamma/N + t_\gamma) - (\alpha + \mu) I_\gamma^T \end{array}\right. \tag{1}$$

with initial conditions

$$\begin{array}{l} S(0) = S_0, E_s(0) = E_0^s, E_s^T(0) = E_{0T}^s, \\ E_\gamma(0) = E_0^\gamma, E_\gamma^T(0) = E_{0T}^\gamma, I_s(0) = I_0^s, \\ I_s^T(0) = I_{0T}^s, I_\gamma(0) = I_0^\gamma, I_\gamma^T(0) = I_{0T}^\gamma. \end{array} \tag{2}$$

Table 1. Variables and parameters of SEIS model of Tuberculosis spread.

Symbol	Description	Value	Ref.
Variables (ppl.)			
S	susceptibles and carries of latent infection		
E_s, E_s^T	TB infected without bacterioexcretion (TB BE-)		
E_γ, E_γ^T	MDR-TB infected without bacterioexcretion (MDR-TB BE-)		
I_s, I_s^T	TB infected with bacterioexcretion (TB BE+)		
I_γ, I_γ^T	MDR-TB infected with bacterioexcretion (MDR-TB BE+)		
Main parameters (parameters of inflow)			
b	birth rate	see Sec. 3	[42]
β	contagiousness of TB contact with bacterioexcretion	[0, 10]	[39]
k	population potential (ppl.)	$4 \cdot 10^6$	
t_γ	fraction of non-MDR infected turning into MDR infected per year	[0.001, 0.1]	f.33
ω	rate of TB BE- turning into BE+ per year	[0.1, 4]	[39]
Outflow parameters			
μ	speed of population outflow	see Sec. 3	[42]
α	additional death rate of TB infected with bacterioexcretion	[0.03, 0.2]	f.33
Treatment parameters			
φ^-	rate of undetected TB BE- infected, that is detected per year	[0.2, 0.5]	f.33
φ^+	rate of undetected TB BE+ infected, that is detected per year	[0.5, 3]	f.33
γ_1	rate of detected MDR-TB infected, that is treated per year	[0.25, 0.5]	f.33
γ_2	rate of detected TB infected, that is treated per year	[0.3, 2]	f.33
Initial state			
σ^-	fraction of undetected TB BE- infected in the initial moment	[0, 0.5]	
σ^+	fraction of undetected TB BE+ infected in the initial moment	[0, 0.5]	

Here $N = S(t) + E_s(t) + E_s^T(t) + E_\gamma(t) + E_\gamma^T(t) + I_s(t) + I_s^T(t) + I_\gamma(t) + I_\gamma^T(t)$ is a volume of the studied population that satisfies the logistic equation:

$$\frac{dN}{dt} = (b - \mu) \left(1 - \frac{N}{k}\right) N - \alpha(I_s + I_\gamma + I_s^T + I_\gamma^T).$$

Here $b - \mu > 0$, $bN - mN^2$ refers to the influx of susceptible individuals into the population due to the birth rate or immigration, $m = (b - \mu)/k$, k is a population potential.

The population is divided into nine non-overlapping groups according to Figure 2 (the subscript T denotes the group of detected infected individuals). The description of the variables and parameters of the model is given in Table 1.

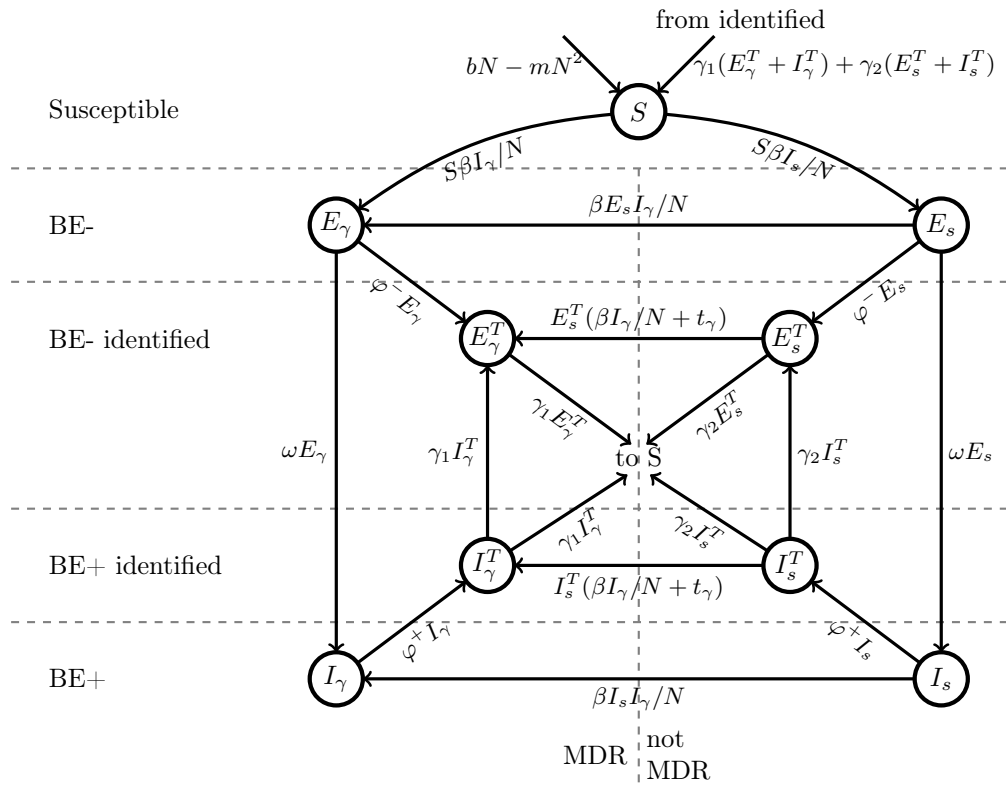


Figure 2. Schematic diagram of the Tuberculosis dynamics model with MDR (1), based on the mass conservation law.

This model is based on the following principles for Russian Federation:

- TB BE- and TB BE+ forms are distinguished, and TB BE- is considered an early stage of the disease preceding the development of severe TB BE+ stages;
- We assume treated individuals return to full susceptibility S , consistent with evidence of high reinfection rates in high-incidence settings [43,44];
- patients are divided into detected and undetected;
- the proportion of those cured from TB with bacterial excretion can transit to a form without bacterial excretion;
- the probability of transition from forms without MDR to TB-MDR is nonlinear;
- during treatment of TB BE+ infected people can transit to both latent forms S and TB BE- with equal probability.

A part of TB patients who successfully complete treatment, when the symptoms have disappeared and bacterial excretion has stably stopped, undergo recurrence. In the Siberian Federal District in 2024 the recurrence rate among those cured averages 33% (in different regions – from 20 to 43% [45]). This means that cured patients remain susceptible to developing TB infection. TB recurrence results from immunological deficiency (dependent on risk factors such as unfavorable social, medical and stress factors, etc.) and is caused by either relapse of an original infection or exogenous reinfection with a new or the same strain of *M. tuberculosis*. The source of relapse is bacteria that remain in a persistent, viable state in cured (BE-) persons for many months and years within post-TB residual lung foci, lymph

nodes, or even within hematopoietic or mesenchymal stem cells [46–49]. The source of reinfection is bacteria acquired through continued contacts of cured patients with people with subclinical or active TB [32,44]. In addition, indirect evidence that cured patients remain in a TB-sensitive risk group is that IGRA tests after cure remain positive in the absolute majority of cases for decades and even until the end of life [50].

Within the framework of the model, all populations, regardless of status, are subject to average mortality rate with the parameter μ . A susceptible or latent infected individual from group S moves to the infected non-bacterial stage E_s (or closed form of TB) after contact with an TB infected without MDR I_s governing the contagiousness parameter β or to the TB infected non-bacterial stage with MDR E_γ . Infected closed-form TB E_s, E_γ migrate to the bacterial form with ω probability. Infected patients without MDR can migrate to infected MDR-TB with t_γ probability, and also with rate β when in contact with infected TB without bacterioexcretion. In the model, each group of TB infected people is divided into detected (with index T) and undetected with expression coefficients φ^- for TB BE-, MDR-TB BE- and φ^+ for TB BE+, MDR-TB BE+, respectively. In the model, additional mortality of infected people with bacterioexcretion α is observed and they move to the state S with probabilities γ_1 (for form with MDR) and γ_2 .

The model principle construction was based on real infection reserwars. This is TB with bacterioexcretion and TB without bacterioexcretion - the incipient and subclinical TB, the latter always bearing a real hazard of converting to bacterioexcretion clinical form. That does not mean, that we exclude significance of latent TB, especially in HIV-infected individuals, but consider this of lower value and as perspective of further studies with the focci of the ending TB.

The limits of the change in the parameters of the models, that are used in linear parts of equations, were determined from statistics (TB-infected registration form No. 33). For example, the proportion of identified non-MDR TB cured in 1 year, γ_2 in a linear approximation is inversely proportional to the time spent in therapy. That is, the limits $\gamma_2 \in [0.3, 2]$ correspond to $[0.5, 3.33]$ years undergoing therapy for TB+ non-MDR patients.

Restrictions on initial conditions in the regions of the Russian Federation are also assumed, namely the percentage of undetected infected people is less than 50% due to the absence of a sharp outbreak of TB:

$$\begin{aligned} E_0^s &= \sigma^- E_{0T}^s, E_0^\gamma = \sigma^- E_{0T}^\gamma, \\ I_0^s &= \sigma^+ I_{0T}^s, I_0^\gamma = \sigma^+ I_{0T}^\gamma. \end{aligned}$$

2.1. ODE System Analysis

The behavior of the ODE systems is broadly studied from the perspective of asymptotic properties, and this paper is also the case. We note that for such systems it is often possible to find analytical solution using Lie algebra [51], however it is effectively applicable only for smaller systems, or systems with symmetries. Thus, the model considered is too complicated for this approach, so we present only the results of stationary points stability.

For the derivation of stationary points of the system (1), only in this section we assume that $\gamma_1 = \gamma_2 = \gamma$. Let us consider the system obtained from the system (1) by adding the equations and replacing:

$$\begin{aligned} I &= I_s + I_\gamma, I^T = I_s^T + I_\gamma^T, \\ E &= E_s + E_\gamma, E^T = E_s^T + E_\gamma^T. \end{aligned}$$

Then we obtain the following system:

$$\begin{cases} \frac{dS}{dt} &= bN - mN^2 - \beta SI/N + \gamma I^T - \mu S; \\ \frac{dE}{dt} &= \beta SI/N - (\mu + \omega + \varphi^-)E; \\ \frac{dE^T}{dt} &= \varphi^- E + \gamma I^T - \mu E^T; \\ \frac{dI}{dt} &= \omega E - (\alpha + \mu + \varphi^+)I; \\ \frac{dI^T}{dt} &= \varphi^+ I - (2\gamma + \alpha + \mu)I^T. \end{cases} \quad (3)$$

Lemma 1. *The system (3) has the only stationary state that is trivial.*

Proof. By equating the right-hand sides to 0, we obtain a system for finding stationary solutions:

$$\begin{cases} \beta SI/N + \mu S &= bN - mN^2 + \gamma I^T; \\ (\mu + \omega + \varphi^-)E &= \beta SI/N; \\ \mu E^T + \gamma I^T &= \varphi^- E; \\ (\alpha + \mu + \varphi^+)I &= \omega E; \\ (2\gamma + \alpha + \mu)I^T &= \varphi^+ I. \end{cases} \quad (4)$$

From 2 and 4 equations of system 4, we obtain:

$$S = \frac{(\mu + \omega + \varphi^-)(\alpha + \mu + \varphi^+)}{\beta\omega} N.$$

Therefore, we are left with equations 1, 3 and 5 of system (4) with a single nonlinear term N^2 .

Note that $N(0) < k$, then $\forall t : N(t) < k$.

In the edge cases, when $N \ll k$ or $N \approx k$ we obtain linear system for stationary point, thus, it is unique (except for the trivial point). As soon as the solution is confined, it may tend towards one of these solutions.

We rewrite the system (4) the following way:

$$\begin{cases} (\mu + \omega + \varphi^-)E + \mu S = bN - mN^2 + \gamma I^T; \\ E^T = I(\varphi^-(\alpha + \mu + \varphi^+)/\omega - \gamma\varphi^+/(2\gamma + \alpha + \mu))/\mu &= Ic_{Et}; \\ E = I(\alpha + \mu + \varphi^+)/\omega &= Ic_E; \\ I^T = I\varphi^+/(2\gamma + \alpha + \mu) &= Ic_{It}. \end{cases}$$

As soon as $dN/dt: bN - mN^2 = \alpha(I + I^T)$, and for the first equation, we get:

$$(\mu + \omega + \varphi^-)Ic_E + \mu S = \alpha I(1 + c_{It}) + \gamma Ic_{It}.$$

Thus,

$$I = -\frac{\mu S}{(\mu + \omega + \varphi^-)c_E - \alpha(1 + c_{It}) - \gamma c_{It}} = c_I S.$$

Then, as soon as $dN/dt = 0$, we get:

$$(b - \mu) \left(1 - \frac{N}{k}\right) N - \alpha(I + I^T) = 0$$

Thus,

$$N = k \left(1 - \frac{\alpha c_S c_I (1 + c_{It})}{b - \mu}\right). \quad (5)$$

That means that there is unique stationary point, except for the trivial and disease-free. However, by substitution of coefficients c_S, c_I, c_E, c_{It} into (5), we obtain $N = 0$ and this point coincides with the trivial point (checked with Wolfram Mathematica, the code is available in GitHub repository). $\square$

Our model operates in absolutes and reflects the logistic curve of the population growth. Most human populations in the regions of the Russian Federation show slow population dynamics and in these dynamics the nontrivial stationary state of TB prevalence can exist. However, we can apply the model to the risk groups where fast growth and diminution can occur.

2.1.1. Stationary Points Stability

The matrix of linearized system (4):

$$\begin{pmatrix} b - 2mN - \frac{\beta I(N-S)}{N^2} - \mu & b - 2mN & b - 2mN & b - 2mN - \frac{\beta S(N-I)}{N^2} & b - 2mN + \gamma \\ \frac{\beta I(N-S)}{N^2} & -\mu - \omega - \varphi^- & 0 & \frac{\beta S(N-I)}{N^2} & 0 \\ 0 & \varphi^- & -\mu & 0 & \gamma \\ 0 & \omega & 0 & -\alpha - \mu - \varphi^+ & 0 \\ 0 & 0 & 0 & \varphi^+ & -2\gamma - \alpha - \mu \end{pmatrix} \quad (6)$$

There are two stationary points $N = 0$ and $S = N = k$.

Then from the form of (6) with $I = 0$, point $S = N = k$ is stable, when

$$\mu < b \quad \text{and} \quad (\mu + \omega + \varphi^-)(\alpha + \mu + \varphi^+) < \beta\omega \quad (7)$$

and point $N = 0$ is stable when $b - \mu < 0$, or $\mu > b$.

In our examples we will consider cases $b < \mu$ and $b \approx \mu$, what depicts the situation in Russian Federation [52]. That means that the point $N = 0$ is globally stable and the stability analysis of stationary points is not informative, since the difference $\mu - b$ is quite small and thus the convergence to the point $N = 0$ is slow, so we are interested in short-term dynamics (compared to the convergence to stationary point).

The 10-year window may seem too large for short-term forecasting; however, that is not the case for Tuberculosis and the current difference between birth and mortality rates. The Tuberculosis has the infectivity rate around 2-4 and recovery rate around 0.1 (for annual time scale). It indicates that the Tuberculosis infection characteristic duration time is around 10 years. The difference between birth and mortality rate is too small (0.003), that implies that the solution will reach the point with $N = 0$ in more than 4000 years if the rates stay constant.

2.1.2. Basic Reproduction Number $\mathcal{R}_0$

To derive the basic reproduction number $\mathcal{R}_0$, we consider two states of people in the system (1): infected without MDR $x_1 = (E_s, E_s^T, I_s, I_s^T)^T$ and with MDR-TB $x_2 = (E_\gamma, E_\gamma^T, I_\gamma, I_\gamma^T)^T$ at the disease free equilibrium (DFE) point $\mathcal{E}_0 = (k, 0, 0, 0, 0, 0, 0, 0)$, i.e. $S_0 = N = k$. Following the approach from the article [53], we derive the basic reproduction number for non-MDR TB $\mathcal{R}_1$ and MDR-TB $\mathcal{R}_2$.

We rewrite the system (1) as follows:

$$\dot{x}_i = (f_i - v_i)x_i, \quad i = 1, 2,$$

where f_i characterize infection in the group due to external flows, v_i – recovery, mortality in the group. Then we have:

$$f_1 = \begin{pmatrix} \beta S I_s / N \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad v_1 = \begin{pmatrix} \frac{\beta E_s I_\gamma}{N} + \omega E_s + \varphi^- E_s + \mu E_s \\ \frac{\beta E_s^T I_\gamma}{N} + \gamma_2 E_s^T - \gamma_2 I_s^T - \varphi^- E_s + t_\gamma E_s^T + \mu E_s^T \\ \frac{\beta I_s I_\gamma}{N} - \omega E_s + \varphi^+ I_s + (\alpha + \mu) I_s \\ \frac{\beta I_s^T I_\gamma}{N} + 2\gamma_2 I_s^T - \varphi^+ I_s + t_\gamma I_s^T + (\alpha + \mu) I_s^T \end{pmatrix}.$$

$$f_2 = \begin{pmatrix} \frac{\beta S I_\gamma}{N} + \frac{\beta E_s I_\gamma}{N} \\ \frac{\beta E_s^T I_\gamma}{N} \\ \frac{\beta I_s I_\gamma}{N} \\ \frac{\beta I_s^T I_\gamma}{N} \end{pmatrix}, \quad v_2 = \begin{pmatrix} \omega E_\gamma + \varphi^- E_\gamma + \mu E_\gamma \\ \gamma_1 E_\gamma^T - \gamma_1 I_\gamma^T - \varphi^- E_\gamma - t_\gamma E_s^T + \mu E_\gamma^T \\ -\omega E_\gamma + \varphi^+ I_\gamma + (\alpha + \mu) I_\gamma \\ 2\gamma_1 I_\gamma^T - \varphi^+ I_\gamma - t_\gamma I_s^T + (\alpha + \mu) I_\gamma^T \end{pmatrix}.$$

For the linearized system $\dot{x}_i = (F_i - V_i)x_i$, $i = 1, 2$, we calculate the Jacobian matrices

$$F_i = \frac{\partial f_k}{\partial x_i^j}(\mathcal{E}_0), \quad V_i = \frac{\partial v_k}{\partial x_i^j}(\mathcal{E}_0), \quad j, k = 1, \dots, 4,$$

after which we find the eigenvalues of the matrices product: $\lambda(F_i V_i^{-1})$, $i = 1, 2$. The matrices F_i and V_i are presented in Appendix A.1.

Thus, we get

$$\mathcal{R}_1 = \mathcal{R}_2 = \frac{\beta\omega}{(\omega + \varphi^- + \mu)(\varphi^+ + \alpha + \mu)}.$$

As soon as $\mathcal{R}_0 = \max\{\mathcal{R}_1, \mathcal{R}_2\}$, the basic reproduction number in the model (1) has the form

$$\mathcal{R}_0 = \frac{\beta\omega}{(\omega + \varphi^- + \mu)(\varphi^+ + \alpha + \mu)}. \quad (8)$$

The basic reproduction number $\mathcal{R}_0$ is based on infectious rates (contagiousness of TB, rate of undetected TB, rate of TB BE- turning into BE+). The difference between the MDR and non-MDR strains in the model (1) is the rate of treatment, which is not reflected in the basic reproduction number (since it describes the rate of infection in a non-immune population). The identical $\mathcal{R}_0$ suggests that, in the absence of treatment, MDR and non-MDR spread similarly. This highlights the role of treatment in controlling MDR.

Lemma. *The stationary point without infection $\mathcal{E}_0$ is stable when basic reproduction number (8) $\mathcal{R}_0 < 1$.*

The $\mathcal{R}_0 < 1$ in Lemma means, that $\beta\omega < (\omega + \varphi^- + \mu)(\varphi^+ + \alpha + \mu)$. It coincides with part of (7), that represents stability with respect to infection outbreak.

The exclusion of latent forms of TB increases the basic reproduction number $\mathcal{R}_0$, namely, due to undetected infectiousness and the transition to an active form (it is possible that such a transition may not occur during the simulation), the $\mathcal{R}_0$ will decrease (see details in Appendix A.1).

2.2. Inverse Problem

We assume that additional information about solution in fixed points in time is given:

$$\begin{aligned} E_s^T(t_j) + E_\gamma^T(t_j) + \varepsilon_j &= E_j^I, \\ I_s^T(t_j) + \varepsilon_j &= I_{s,j}^I, \quad I_\gamma^T(t_j) + \varepsilon_j = I_{\gamma,j}^I, \\ j &= 1, \dots, M. \end{aligned} \quad (9)$$

Then the inverse problem (1), (2), (9) is to identify the parameters of the SEIS model $q = (\beta, \gamma_1, \gamma_2, \omega, t_\gamma, \alpha, \varphi^-, \varphi^+, \sigma^-, \sigma^+) \in \mathbb{R}_+^{10}$ with the help of additional information (9).

We assume that observations $y_t = (E_s^T + E_\gamma^T, I_s^T, I_\gamma^T)$ are subject to independent additive Gaussian errors:

$$y_t = f_t(q) + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma).$$

Here $f_t(q) = (E_s^T(t; q) + E_\gamma^T(t; q), I_s^T(t; q), I_\gamma^T(t; q))$, $\Sigma = \text{diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ and σ_i are estimated jointly with q .

The model (1) is theoretically identifiable based on algebra analysis [54].

Due to the instability and non-uniqueness of the solution to the inverse problem, the following section presents a sensitivity analysis of the model, which will allow us to estimate the degree of sensitivity of the parameters q to the measured states of the system (9) [55].

2.3. Sensitivity-Based Identifiability Analysis for SEIS Model

The sensitivity analysis of the model (1) was carried out using the Sobol method [56], with a uniform distribution of parameters, the variation limits of which are given in Table 1. This method quantifies how the variance in model outputs (predictions of TB diagnoses) can be apportioned to individual parameters and their interactions across the entire apriori feasible parameter space. The initial conditions (2) approximate the epidemiological situation of the Novosibirsk region in 2009:

$$\begin{aligned} N &= 2263\,383, \quad E_s = 300, \quad E_\gamma = 500, \\ E_s^T &= 1\,100, \quad E_\gamma^T = 1\,549, \\ I_s &= 100, \quad I_\gamma = 400, \quad I_s^T = 438, \quad I_\gamma^T = 1\,330. \end{aligned}$$

Thus, $S = N - E_s - E_\gamma - E_s^T - E_\gamma^T - I_s - I_\gamma - I_s^T - I_\gamma^T = 2257\,666$ people. The modeling time is 10 years.

For a model outputs as we consider in inverse problems $y_1 = E_s^T(t; q) + E_\gamma^T(t; q)$, $y_2 = I_s^T(t; q)$, $y_3 = I_\gamma^T(t; q)$ (see measurement 9) the Sobol decomposition expresses the total variance $V(y_n)$, $n = 1, 2, 3$:

$$V^{(n)} = V(y_n) = \sum_i V_i^{(n)} + \sum_{i < j} V_{ij}^{(n)} + \dots + V_{1\dots m}^{(n)}, \quad n = 1, 2, 3.$$

Here $V_i^{(n)} = V[E(y_n|q_i)]$, $i = 1, \dots, 10$, is the first-order (main) effect variance, measuring the contribution of parameter q_i alone, $V_{ij}^{(n)}$ is the second-order interaction variance between q_i and q_j , and so on for higher-order interactions.

We compute the total-order Sobol index

$$S_{T_i}^{(n)} = \frac{V^{(n)} - V[E(y_n|\mathbf{q}_{\sim i})]}{V^{(n)}}, \quad n = 1, 2, 3,$$

and the first-order and second order Sobol indexes

$$S_{1_i}^{(n)} = \frac{V_i^{(n)}}{V^{(n)}}, \quad S_{ij}^{(n)} = \frac{V_{ij}^{(n)}}{V^{(n)}}, \quad n = 1, 2, 3,$$

where $\mathbf{q}_{\sim i}$ denotes all parameters except q_i , $1 = 1, \dots, 10$. $S_{T_i}^{(n)}$ measures the total contribution of q_i , including all its interactions with other parameters. A parameter with very low total-order indices $S_{T_i}^{(n)}$ (e.g., < 0.01) has a negligible influence on the measurements y_n , $n = 1, 2, 3$. Its value cannot be inferred from the data. If $S_{T_i}^{(n)} > 0.5$ for some q_i to data y_n than total contribution of q_i , including all its interactions with other parameters is significant. However, if $S_{T_i}^{(n)} > 0.5$ and $S_{1_i}^{(n)} < 0.5$ than parameter q_i may still be identifiable, but its estimate will be strongly correlated with others.

We use Sensitivity Analysis Library in Python (SALib package) and set $N = 1024 \cdot (2 \cdot 8 + 2) = 18432$ Sobol samples. In Figure 3 (right) the total-order Sobol coefficients are given for the model coefficients $\tilde{q} = (\beta, \gamma_1, \gamma_2, \omega, t_\gamma, \alpha, \varphi^-, \varphi^+) \in \mathbb{R}_+^8$ (without initial conditions), including the first indices, the second (correlation), etc. with 95% confidence intervals. The most sensitive to measurements of detected infected is the TB contagiousness parameter β . The TB detection parameter BE+ φ^+ is sensitive to infected TB BE+ without MDR $I_s^T(t_j)$. The proportion of infected TB converting to the BE+ form ω is sensitive to infected TB-MDR BE+ with and without MDR. The additional mortality rate α due to TB is sensitive to MDR-TB BE+ measurements.

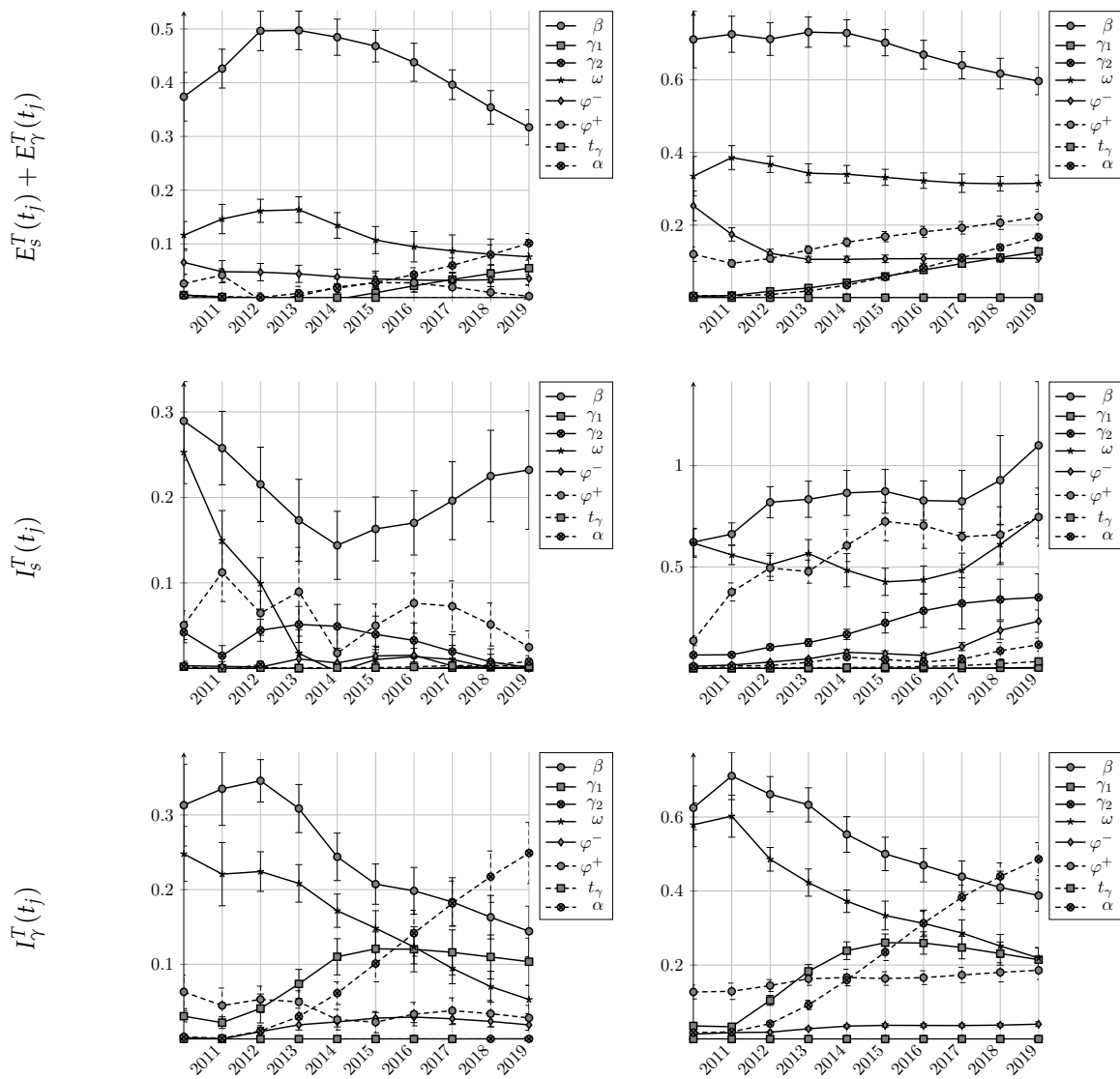


Figure 3. The first Sobol indices $S_{1_i}^{(n)}$ (left) and the total Sobol indices $S_{T_i}^{(n)}$ (right) for parameters $\tilde{q}_i$, $i = 1, \dots, 8$, with 95% confidence intervals for 10 years of simulation for TB BE- (the first row), TB BE+ without MDR $I_s^T(t)$ (the middle row) and MDR-TB BE+ $I_\gamma^T(t)$ (the last row).

Note, that the first Sobol index for β coefficient $S_1^{(1)}$ for TB BE- measurement is more than 0.5 (this parameter is sensitive to the y_1), other ones are less than 0.5 that means the parameter’s influence is primarily through interactions (see Figure 3 left). Its may still be identifiable, but its estimate will be strongly correlated with others. The second Sobol indexes $S_{\beta\omega}^{(2)}, S_{\beta\omega}^{(3)}, S_{\beta\varphi^+}^{(2)}$ more than 0.2 and $S_{\beta j}^{(2)}$ close to 0.1-0.2 for all parameters j (see correlation matrices in the Appendix A.2).

As we note that Sobol indices show strong parameter correlations, the formal practical identifiability analysis is performed, i.e. we construct the correlation matrix (see Table 2 for Novosibirsk region) based on the Fisher information matrix

$$\rho_{ij} = \frac{FIM_{ij}^{-1}}{\sqrt{FIM_{ii}^{-1}FIM_{jj}^{-1}}},$$

where

$$FIM = \left(\frac{\partial f_t}{\partial q}\right)^T \Sigma^{-1} \frac{\partial f_t}{\partial q}$$

and $\frac{\partial f_t}{\partial q}$ is a sensitivity matrix.

Table 2. Correlation matrix ρ_{ij} for Novosibirsk posterior.

	β	γ_1	γ_2	ω	φ^-	φ^+	t_γ	α
β	1	-0.609	-0.739	0.309	-0.886	0.984	0.881	0.869
γ_1	-0.609	1	0.947	-0.481	0.875	-0.506	-0.669	-0.849
γ_2	-0.739	0.947	1	-0.559	0.926	-0.639	-0.855	-0.935
ω	0.309	-0.481	-0.559	1	-0.294	0.145	0.644	0.666
φ^-	-0.886	0.875	0.926	-0.294	1	-0.843	-0.825	-0.901
φ^+	0.984	-0.506	-0.639	0.145	-0.843	1	0.800	0.771
t_γ	0.881	-0.669	-0.855	0.644	-0.825	0.800	1	0.943
α	0.869	-0.849	-0.935	0.666	-0.901	0.771	0.943	1

Table 2 shows that the parameters with $|\rho_{ij}| > 0.95$ are difficult to distinguish. For example, $\rho_{\beta\varphi^+} > 0.98$ indicates that the data cannot independently inform the parameters β and φ^+ , only their product is well-identified. That are shown during the sensitivity Sobol analysis as well.

3. Data Analysis

The birth b and mortality μ rates are presented in Tables 3 and 4 for five regions of Russian Federation.

Table 3. Number of deaths in Russian Federation regions per 1000 residents.

Region	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Zabaikalskiy krai	13.8	13.8	13.3	13.1	12.5	12.5	12.9	12.3	11.7	12.3	12.4	13.7	15.8	13.8
Tyva republic	12	11.6	11	11.2	10.9	10.9	10.3	9.8	8.7	8.8	8.3	9.4	9	8.6
Novosibirsk oblast	14	13.9	13.6	13.6	13.4	13.3	13.1	13	12.9	12.9	12.7	15.3	17	13.7
Irkuts oblast	14.3	14.4	14	13.9	13.6	13.7	13.6	13.4	12.9	13	13.2	15	17.7	14.1
Altai republic	14.7	15	14.6	14.6	14.2	14.2	14.1	14.1	14	14.3	14	16.5	19.1	15.8

Table 4. Number of births in Russian Federation regions per 1000 residents.

Region	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Zabaikalskiy krai	16.1	15.9	15.6	16.3	16.1	16.2	15.7	14.9	13.7	13	12.2	12.2	11.9	11.2
Tyva republic	26.9	26.8	27.4	26.6	26	25.2	23.7	23.1	21.8	20.1	18.4	20	19.7	17.7
Novosibirsk oblast	12.9	13.2	13.1	13.9	14.1	14	14.2	13.7	12.3	11.7	10.7	10.3	10.1	9.6
Irkuts oblast	15.6	15.2	15.3	15.9	15.7	15.2	15.3	14.7	13.4	12.8	11.8	11.3	11	10.4
Altai republic	12.7	12.7	12.8	13.8	13.6	13.4	12.9	12.4	11.2	10.3	9.4	9	8.7	8.2

To restore parameters of the model for regions of Russian Federation, we used the following data:

- Let $p_\gamma = I_\gamma^T / (I_\gamma^T + I_s^T)$ then from [43], we have data of fraction and absolute number of patients with MDR among the patients with bacillary forms of TB, that represent p_γ and I_γ^T respectively.
- From data of form 33 (Information about patients with Tuberculosis from the Order of Rosstat dated 31.12.2010 №483), we have cohorts that represent $I_s^T + I_\gamma^T + E_s^T + E_\gamma^T = C_{EI}$. Then we may obtain the following:
 - I_γ^T is given in data;
 - $I_s^T = I_\gamma^T (1/p_\gamma - 1)$;
 - $E_\gamma^T + E_s^T = C_{ET} - I_\gamma^T - I_s^T$.

For demonstration of the numerical results we choose two regions of Russian Federation with an unfavorable situation in terms of the prevalence of TB: Novosibirsk region and Tyva Republic.

4. Model Scenarios of Tuberculosis Spread

As a result of solving the inverse problem (1), (2), (9), the vector of SEIS model parameters q was determined for each region, the substitution of which into the model (1)-(2) leads to a comparison of the model and real data on identified infected people (Figure 5).

4.1. Bayesian Approach

To obtain the distribution of the model parameters and to construct a forecast considering the uncertainty, we use the Bayesian MCMC (Markov Chain Monte Carlo) approach, in which the posterior distribution for the parameters $q \in \mathbb{R}_+^{10}$ has the form:

$$p(q | y) \propto p(y | q) \cdot \prod_{i=1}^{10} p(q_i).$$

Here $p(q_i)$ is the prior distributions of the SEIS model parameters, $p(y | q)$ is the likelihood of the data.

The likelihood functional is defined by the logarithmic density function of the normal distribution with the mean being the difference of the real data and solution of the direct problem for the given current parameters. The variance of likelihood functional is determined during the MCMC algorithm (see Appendix A.3).

$$\beta \ t_\gamma \ \omega \ \alpha \ \varphi^- \ \varphi^+ \ \gamma_1 \ \gamma_2 \ \varphi^- \ \varphi^+$$

The results of the inverse problem solution are in Table 5. Parameters with total Sobol index less than 0.01 for all outputs were considered practically non-identifiable and assigned informative Normal priors centered at Optuna values with width 10% of the prior range (see results for mean values in Appendix A.4). The prior distribution $p(q_i)$ for sensitive parameters β, φ^+, ω is proposed as uniform (see Table 1).

We apply the Bayesian approach to find the posteriority distribution of the parameters q with 20000 draws of 4 Markov chains (i.e. 80000 samples for each parameter), mean (column "Mean") and 95% credible interval (CI). Differences in the parameters of pathogen contagiousness β , the proportion of infected TB converting to the BE+ form ω and the rate of detection φ^+ of TB infected individuals with bacterial excretion indicate heterogeneity in the prevalence of TB in the Russian Federation regions.

The Pareto front (see Figure A3, Appendix A.5) for three measured statistics (9) shows the uniqueness of the given solution of the inverse problem (TB parameter identification), then the CIs values depend on numerical accuracy the of the correlated parameters.

Table 5. Restored parameters of SEIS model of Tuberculosis spread (1) in regions of Russian Federation.

Symbol	A priori interval	Novosibirsk		Tyva		Altai		Irkutsk		Zabaikal	
		Mean	95% CI	Mean	95% CI	Mean	95% CI	Mean	95% CI	Mean	95% CI
R_0	-	0,973	[0,594; 1,598]	0,974	[0,683; 1,400]	0,942	[0,557; 1,595]	0,876	[0,518; 1,502]	0,926	[0,540; 1,593]
β	[0.03, 0.2]	2.676	[2.369; 2.928]	2.972	[2.718; 3.171]	3.566	[3.199; 3.909]	2.927	[2.628; 3.228]	2.128	[1.876; 2.375]
t_γ	[0, 10]	0.057	[0.046; 0.068]	0.022	[0.018; 0.026]	0.019	[0.015; 0.023]	0.049	[0.039; 0.058]	0.068	[0.056; 0.081]
ω	[0.25, 0.5]	3.164	[2.727; 3.588]	3.725	[3.272; 3.983]	1.368	[1.143; 1.597]	1.605	[1.348; 1.871]	0.844	[0.725; 0.971]
α	[0.3, 2]	0.065	[0.053; 0.077]	0.038	[0.031; 0.045]	0.150	[0.120; 0.181]	0.066	[0.054; 0.079]	0.125	[0.101; 0.149]
φ^-	[0, 0.5]	0.326	[0.265; 0.391]	0.267	[0.212; 0.321]	0.488	[0.422; 0.556]	0.295	[0.243; 0.350]	0.275	[0.235; 0.318]
φ^+	[0, 0.5]	2.407	[2.123; 2.634]	2.787	[2.536; 2.957]	2.610	[2.349; 2.836]	2.724	[2.442; 2.968]	1.577	[1.376; 1.772]
γ_1	[0.1, 4]	0.373	[0.326; 0.419]	0.290	[0.238; 0.342]	0.328	[0.273; 0.390]	0.307	[0.258; 0.356]	0.286	[0.245; 0.326]
γ_2	[0.2, 0.5]	0.519	[0.455; 0.585]	0.870	[0.741; 1.012]	0.674	[0.552; 0.816]	0.275	[0.227; 0.323]	0.292	[0.249; 0.338]
φ^-	[0.5, 3]	0.278	[0.230; 0.322]	0.186	[0.152; 0.218]	0.482	[0.447; 0.498]	0.254	[0.213; 0.295]	0.281	[0.234; 0.326]
φ^+	[0.001, 0.1]	0.354	[0.298; 0.411]	0.381	[0.321; 0.440]	0.476	[0.429; 0.499]	0.376	[0.319; 0.433]	0.471	[0.420; 0.497]

Figure 4 demonstrates the mean and CI of model parameters as noted in Table 5. Note, that all parameters β, ω and φ^+ with wide CI are sensitive to the measurement with its interaction to each other (see Section 2.3), i.e. the total-order Sobol indexes are greater than 0.5, but the first-order indexes are sufficiently small. Scenarios of number of TB BE+ infectious may be unreliable because of I_s^T and I_γ^T depend on the combination of all these parameters. Small errors in identifying any one of them lead to unrealistic forecasts. Therefore, when analyzing forecast scenarios for TB BE+ groups, it is appropriate to use interval estimates rather than point estimates, as was done in this study. While the model is structurally identifiable, strong parameter correlations (seen in Sobol indices) affect practical identifiability.

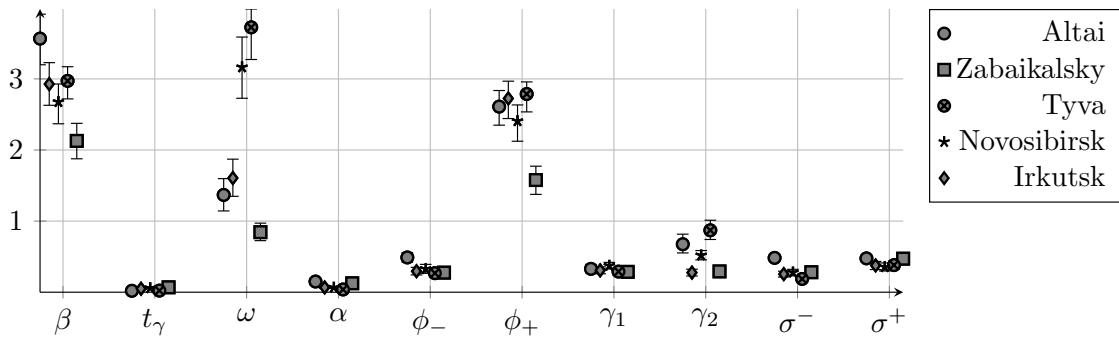


Figure 4. Restored parameters of SEIS model of Tuberculosis spread (1) in regions of Russian Federation with estimations by TPE Optuna optimizer.

In Figure 5–7 the probability distribution of TB infected based on Bayesian approach is demonstrated for Novosibirsk region and Tyva Republic of Russian Federation (Figure 5), Altai and Irkutsk regions (Figure 6) as well as Zabaikalskiy region (Figure 7). The approximation for Novosibirsk region is more accurate because of small 95% credible intervals that contain the real data (dashed dot line). The average for all five regions Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) are in Table 6. For example, for data I_s^T RMSE= $5 \cdot 10^{-5}$ means that the error is 5 per 100000 people. The single situation when mean prediction is worse is the modeling results for I_s^T , what is compensated with adjustments in other groups. The I_s^T forecasts are less accurate only for the Zabaikalsky region, and the data is poorly fit for the Tyva Republic. The Zabaikalsky region data has two “falls” in data, that cannot be described with the chosen model, as it is able to describe only one peak of infection with descent without additional “fall”. The Tyva region data has unusual feature of I_s^T being significantly lower than I_γ^T . This may be due to a simplification of the model by ignoring latent tuberculosis infections, which may actually account for some of the detected cases. The basic reproduction number $\mathcal{R}_0$, calculated using the formula (8), for all considered regions are close to the 0.9.

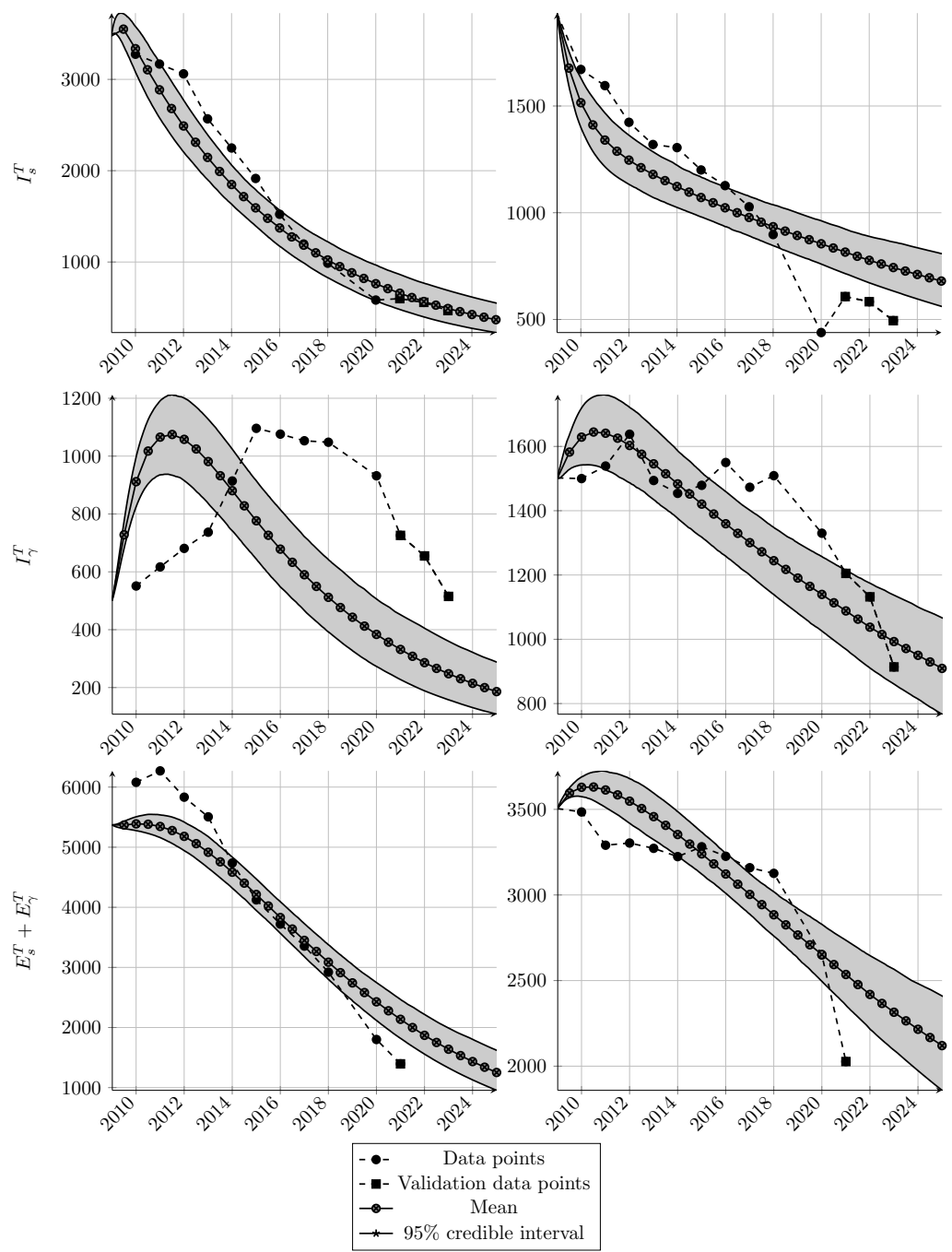


Figure 5. Probabilistic forecast of infected individuals in the Irkutsk region (left) and the Novosibirsk region (right), compared with actual data (dash-dot lines).

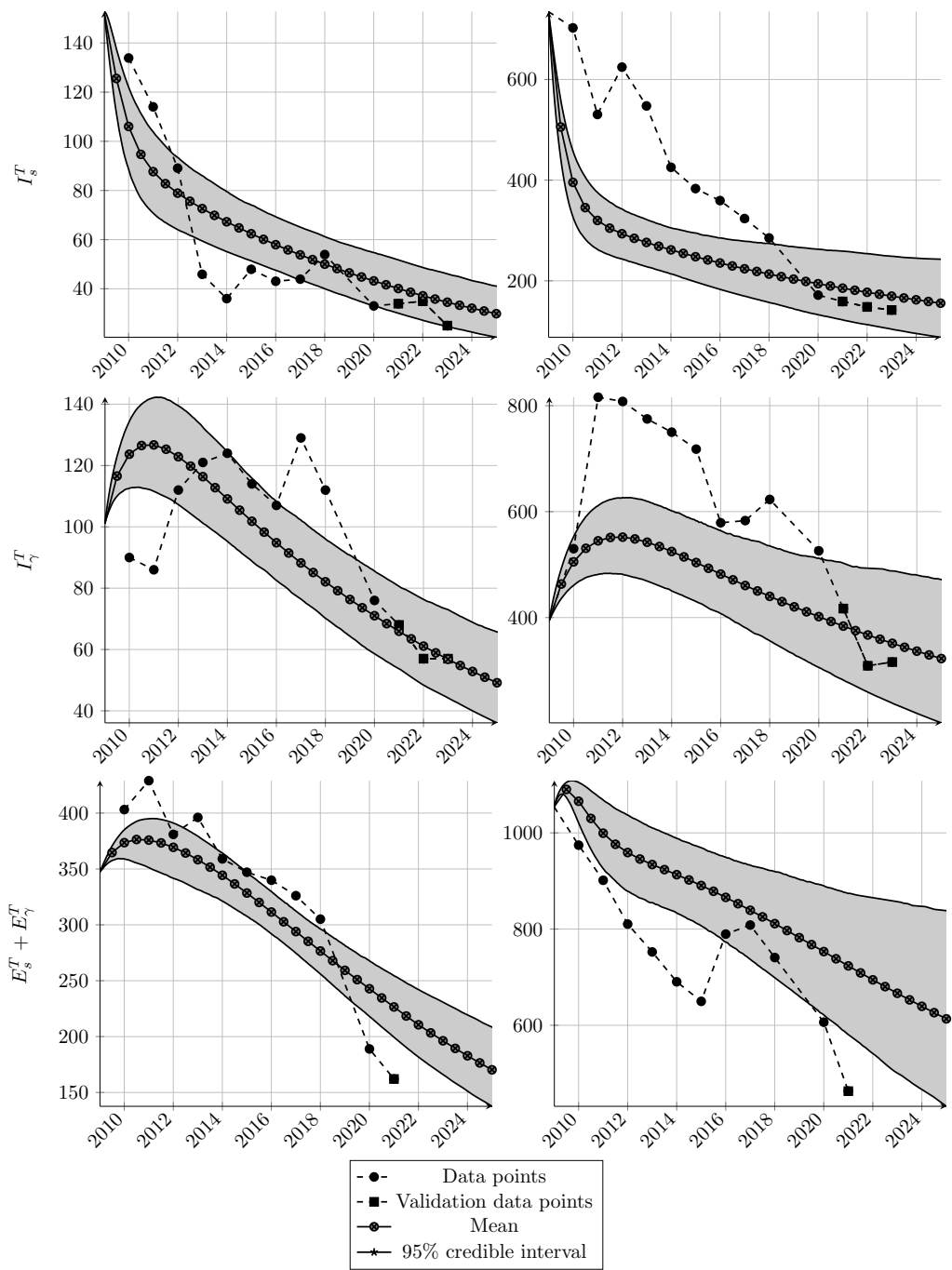


Figure 6. Probabilistic forecast of infected individuals in the Altai republic (left) and the Tyva republic (right), compared with actual data (dash-dot lines).

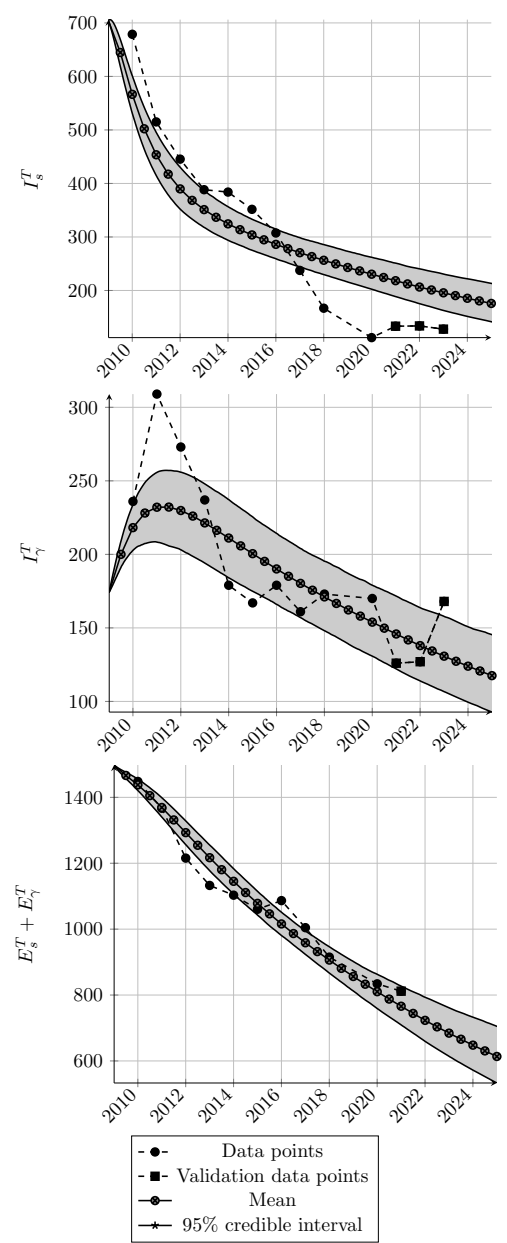


Figure 7. Probabilistic forecast of infected individuals in the Zabaikalsky krai, compared with actual data (dash-dot lines).

Table 6. A quantitative forecast accuracy metrics (MAPE and RMSE) for measured data.

	$E_s^T + E_\gamma^T$	I_s^T	I_γ^T
MAPE	0.2572	0.285	0.2476
RMSE	$2.7 \cdot 10^{-4}$	$5 \cdot 10^{-5}$	$8.7 \cdot 10^{-5}$

We also point out that the data points of 2019 year are not used in inverse problem, because of change in rules of statistical records of TB in Russia [57], which made figures of 2019 year unrepresentative. However, that does not affect the epidemiological process, thus the following years are not affected significantly (see Appendix A.6).

In 2018-early 2019, a surge in TB detection have occurred in the Russian Federation due to the emergence of regulatory official orders that significantly expand the coverage of prophylactic medical examination/observation service and the introduction of modern diagnostic and preventive treatment

methods [58–60]. We suggest that from 2020 onwards, the impact of these official acts was significantly reduced due to the emergence of the COVID-19 pandemic and, as a result, a decrease in TB detection [61].

A short-term forecast for 3 years ahead was constructed for TB infected with bacterioexcretion BE+. The credible interval and forecast accuracy are higher for TB BE+ without MDR. To improve the accuracy of the MDR-TB BE+ forecast, it is necessary to use a priori information on treatment, immunology, etc. The model cannot capture sudden spikes; forecasts assume stable epidemiological conditions.

5. Discussion

The literature on mathematical modeling of the early stages of untreated pulmonary tuberculosis—its incipient and subclinical forms—is extremely scarce. We found modeling of the early stages of TB in the works of Avilov et al. [41,62], where two models were used, one of which assumes the infectiousness of the initial stages of the disease, and the second considers all cases to be initially non-infectious, but some of them subsequently become infectious. The authors do not give preference to either model, although they consider the first model more plausible. Our model is based on the latest ideas about the very high contagiousness of the early stages of the TB process [32,63], and combines both models, assuming the infectiousness of tuberculosis cases in the early stages and the possible gradual development of some cases to an infectious state. At the same time, like Avilov et al. [5] we defined the development of infectiousness as the transition of patients from the BE- to the BE+ state, which allows us to predict the hidden incidence and prevalence of TB within the framework of the model and to better understand the real epidemic situation in different regions.

Coming through recent ideas in TB forecasting based on statistical methods and statistically based machine learning [3], in our work we also relied on long-term statistical data of departmental reporting for the territories of the Siberian and Far Eastern Federal Districts, being the most unfavorable districts in the Russian Federation. However we used the SEIS type generalized compartment model with the law of mass action and machine learning was added not only to the statistical analysis but precisely to the mathematical model.

We consider the closest work to our research is the study [64]. This prototype model comprises 5 differential equations, as our model has 9 differential equations. The authors used common methodology to derive the formula of R_0 and calculate the sensitivity of R_0 . However, the methodology of comparing the real and the model data is completely different in our works.

The very recent interesting work is presented in [65]. The sensitivity analysis has shown that major positive influence on R_0 has the per-capita TB transmission rate, rate of bacteria spread to environment by infectious individuals and the droplet-human propagation coefficient. The negative influence on R_0 has the mycobacterium environmental death rate and the transition rate from the low-risk latent TB to high-risk latent TB. Thou in terms of clinical epidemiology the terminology and methodology in our works are quite different, globally we achieved very similar results.

An analysis of the model sensitivity in our model it was shown that the TB contagiousness parameter β , the proportion of infected TB converting to the BE+ form ω and the TB detection parameter BE+ φ^+ are global sensitive to measurements of the TB BE+ I_s^T , and the first two parameters β and ω are sensitive to the MDR-TB BE+ and TB BE-. The sensitivity is also based on correlations.

Here we fully correspond to the work [66], where the simulation results proved that the key to better prevent and control the spread and develop of TB is to improve the detection rate, and the conversion rate from TB to MDR-TB.

On the Bayesian approach we calculated the credible intervals (CIs) of the posterior distribution of 10 model parameters based for additional information on the number of identified patients with TB BE-, TB BE+ without MDR and MDR-TB BE+ from 2009 to 2021. The less sensitive parameters give small CIs for five considered regions that described the TB propagation and detection in RF regions in common. More sensitive and correlated parameters have huge CIs that proved the heterogeneity of

TB propagation in RF regions as well as opens ways for its control. The Pareto front (see Figure A3) for three measured statistics shows the uniqueness of the given solution of the inverse problem (TB parameter identification), then the CIs values depend on numerical accuracy the of the correlated parameters.

Based on identified parameters and its CIs we construct the forecasting scenarios of detected patients with TB BE-, TB BE+ without MDR and MDR-TB BE+ from 2022 to 2026 years for five regions of the Russian Federation, the approach being close to the results [67,68].

6. Conclusion

The aim of our study was to learn the new patterns and laws of TB spreading worldwide on the example of the Russian Federation and to evaluate the combat strategies. In compartment model with revealed and unrevealed cases, incipient and subclinical TB (BE-) and clinical TB (BE+), drug sensitive and drug resistant TB we obtained the formula for the basic reproduction number $\mathcal{R}_0$. The formula comprises TB contagiousness parameter β , rate of TB BE- converting into BE+ ω and the rate of undetected TB BE+ φ^+ . Using the sensitivity analysis we discovered the leading pattern, determining prevalence of TB clinical forms with bacterial excretion, which are most dangerous for humans in the aspect of lethality and disease transmission. The pattern is based on the same parameters as $\mathcal{R}_0$ determinacy, that are TB contagiousness parameter β , rate of TB BE- converting into BE+ ω and the rate of undetected TB BE+ clinical forms φ^+ . By solving the inverse problem we proved the variability of these crucial parameters in high incidence regions of the Russian Federation. This heterogeneity is due to differences of medical and social systems. The TB contagiousness parameter β depends on intensity of contact tracing. The rate of TB BE- converting into BE+ ω is influenced by the healthcare activity – coverage of the population with X-ray film examination. The detection (reveal) indicator of BE+ clinical forms is based on the same parameter (coverage of the population with X-ray films) and also application of sensitive methods of TB diagnostics, including molecular methods. Intensification of these activities will help us to fulfill the strategy of ending TB. By addressing the prior art gaps: absence of incipient and subclinical disease compartment, case detection rate, consideration of transitions between compartments, use of Bayesian approach for the detection of scenarios – we came closer to understanding the TB epidemic the measures for counteraction.

In the future, to plan detailed parameters of possible interventions and their application in specific regions, additional processes should be included, for example, by introducing quantitative trajectories of socio-economic, medical-social, molecular-epidemiological and other indicators connected to the territories. The most complex and convenient model for use in practical healthcare will have to include an assessment of the financial and economic efficiency of possible interventions. We plan consider a fractional-order extension of the mathematical model to capture memory effects in TB progression and solve the uncertainty of forecasting [69,70].

Acknowledgments: The work was performed according to the Government research assignment for Sobolev Institute of Mathematics SB RAS, project FWNF-2024-0002 "Inverse ill-posed problems and machine learning in biological, socio-economic and ecological processes" (Sections 1-2) and by the Russian Science Foundation, project No. 23-71-10068 (Sections 3-5).

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A.

Appendix A.1. Basic Reproduction Number

The Jacobian matrices ($i = 1, 2$)

$$F_i = \frac{\partial f_k}{\partial x_i^j}(\mathcal{E}_0), \quad V_i = \frac{\partial v_k}{\partial x_i^j}(\mathcal{E}_0), \quad j, k = 1, \dots, 4,$$

for computing basic reproduction number $\mathcal{R}_0$ are as follows

$$F_1 = F_2 = \begin{pmatrix} 0 & 0 & \beta & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$
$$V_1 = \begin{pmatrix} \omega + \varphi^- + \mu & 0 & 0 & 0 \\ -\varphi^- & \gamma_2 + t_\gamma + \mu & 0 & 0 \\ -\omega & 0 & \varphi^+ + \alpha + \mu & 0 \\ 0 & 0 & -\varphi^+ & 2\gamma_2 + t_\gamma + \alpha + \mu \end{pmatrix}.$$

Latent TB influence of the basic reproduction number. The basic reproduction numbers $\mathcal{R}_0$ for models with inclusion of TB latent compartment [5,39,65] contain the rate of transmission from TB latent to subclinical form E (for linear transmission rate σ) as follows:

$$\mathcal{R}_0 \approx \frac{\beta\omega\sigma}{(\omega + \mu + \text{Treatment})(\sigma + \mu + \dots)}.$$

Mathematically it closes to $\mathcal{R}_0$ without TB latent form (8) with the factor $\frac{\sigma}{\sigma + \mu + \dots} < 1$ connected to transmission rate σ from TB latent form to subclinical one. The $\mathcal{R}_0$ becomes less because the latent infectiousness can remain in the incubation period for decades, which exceeds the modeling period.

Appendix A.2. Second-Order Sobol Indices

Based on the sensitivity analysis of TB model (1) we get valuable second-order Sobol indices for sensitive parameters (see Section 2.3 for details). Tables A1 and A2 demonstrate the second-order Sobol indices $S_{jk}^{(i)}$, $i = 2, 3$, for measured data y_i ($y_2 = I_s^T$, $y_3 = I_\gamma^T$) and all pairs (j, k) of parameters $q = (\beta, \gamma_1, \gamma_2, \omega, t_\gamma, \alpha, \varphi^-, \varphi^+)$ for some measured time.

Table A1. Second-order Sobol indices $S^{(2)}$ of TB model to TB diagnosed with BE+ I_s^T for 2019 year.

	β	γ_1	γ_2	ω	φ^-	φ^+	t_γ	α
β	1	0.15	0.25	0.18	0.20	0.21	0.16	0.18
γ_1		1	0	0	0	0	0	0
γ_2			1	0	0	0.02	0.02	0
ω				1	0.02	0	0.01	0
φ^-					1	0.01	0.01	0.02
φ^+						1	0.03	0.04
t_γ							1	0
α								1

Table A2. Second-order Sobol indices $S^{(3)}$ of TB model to MDR-TB diagnosed with BE+ I_γ^T for 2009 year.

	β	γ_1	γ_2	ω	φ^-	φ^+	t_γ	α
β	1	0	0	0.20	0	0.01	0	0
γ_1		1	0	0	0	0	0	0
γ_2			1	0	0	0	0	0
ω				1	0	0.02	0	0
φ^-					1	0	0	0
φ^+						1	0.04	0.04
t_γ							1	0
α								1

High interaction second-order Sobol coefficients ($S > 0.2$) affect the identifiability of parameters and the uncertainty of the forecast (see Figure 4-7), especially at the stage of activation of the form with bacterial excretion (ω coefficient) and MDR-TB forms (β coefficient).

Firstly, we found the correlation between intensity of the TB reveal (φ^- , φ^+) and contagiousness β , for MDR TB reveal the correlation being slightly more strong. That means that the less hidden TB causes are revealed; the more is the intensity of transmission and visa versa. Secondary, we found the correlation between TB contagiousness parameter β and the proportion of infected TB converting to the BE+ form ω and to MDR forms t_γ , the first parameter being of slightly more importance. Together conversion to BE+ and MDR are the main challenges of the contemporary epidemic.

Appendix A.3. MCMC Convergence

A necessary condition for the convergence of Markov chains in the Bayesian method is the analysis of the R-hat statistic. It has been empirically established that if $\hat{R} > 1.1$, then either the Markov chains are insufficiently long (strongly dependent on the a priori assumption) or the chains are "stuck" in certain regions of the parameter space and cannot effectively explore it in the case of complex models and high parameter correlations. The achieved R-hat values for Altai region are displayed in Table A3. All R-hat are less than 1.1 that indicating convergence of MCMC method.

Table A3. R-hat values for parameters of SEIS model of Tuberculosis spread.

Parameter	Irkutsk	Novosibirsk	Tyva	Zabaikal	Altay
β	1.054	1.026	1.027	1.011	1.030
t_γ	1.004	1.017	1.010	1.013	1.005
ω	1.014	1.021	1.010	1.004	1.016
α	1.006	1.009	1.007	1.009	1.006
φ^-	1.016	1.010	1.001	1.004	1.013
φ^+	1.100	1.036	1.035	1.016	1.055
γ_1	1.015	1.021	1.007	1.003	1.006
γ_2	1.005	1.010	1.004	1.009	1.012
σ^-	1.014	1.009	1.005	1.013	1.070
σ^+	1.008	1.019	1.018	1.027	1.045

The convergence process for Monte Carlo trajectories for Altai region is shown in Figure A1.

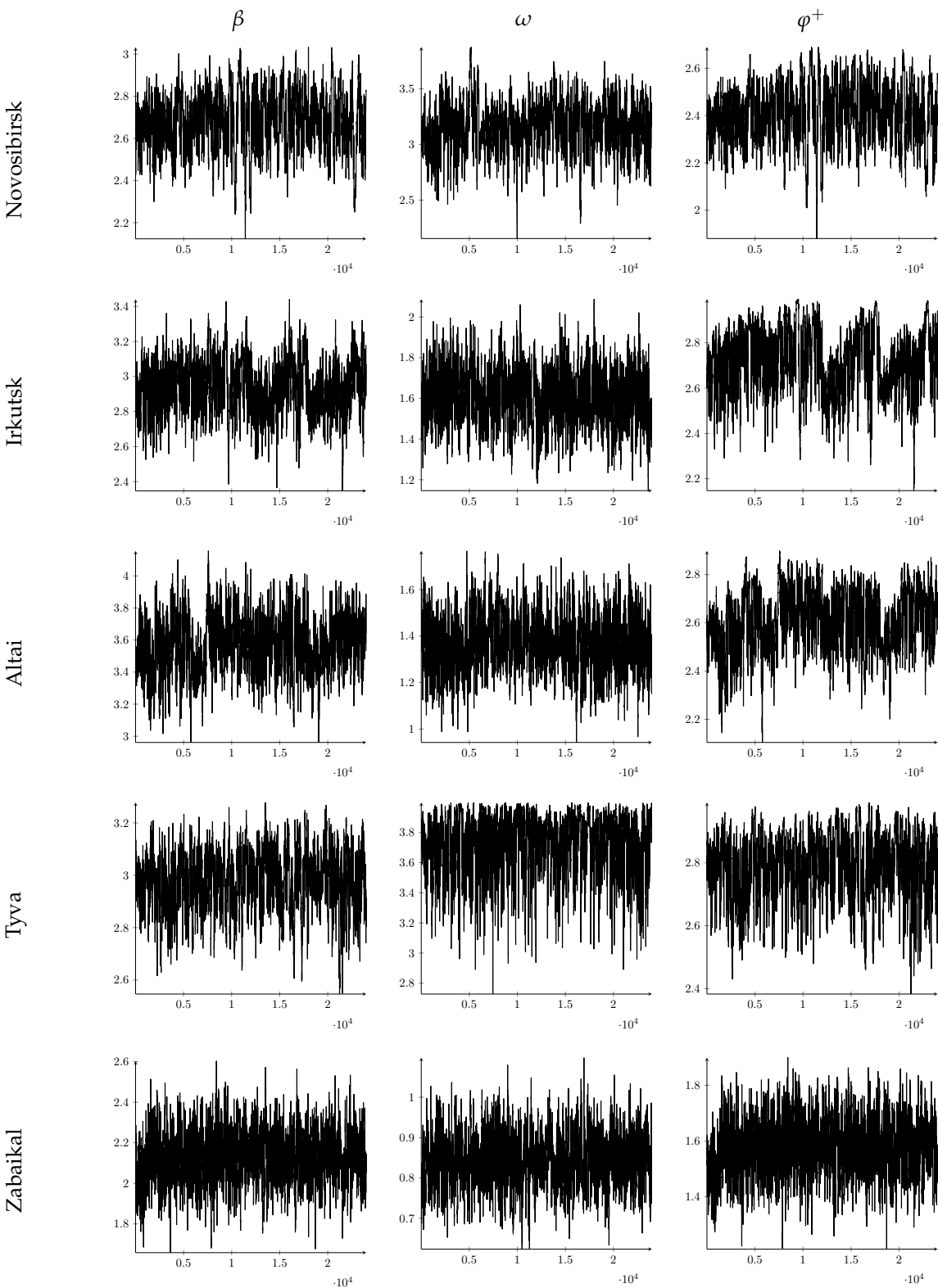


Figure A1. Monte Carlo trajectories for posterior parameters distribution.

Appendix A.4. Inverse Problem: Numerical Approach

The inverse problem (1), (2), (9) may be reduced to minimization problem for the following cost functional:

$$J(q) = \sum_{j=1}^M \frac{(E_s^T(t_j; q) + E_\gamma^T(t_j; q) - E^j)^2}{(E^j)^2} + \frac{(I_s^T(t_j; q) - I_s^j)^2}{(I_s^j)^2} + \frac{(I_\gamma^T(t_j; q) - I_\gamma^j)^2}{(I_\gamma^j)^2}. \quad (A1)$$

We note that this form of cost functional makes us implicitly assume that noise in data is distributed normally.

As a method for solving the problem of minimizing the functional (A1), we use the tree-based Parzen optimization approach included in the Optuna software package implemented in Python (the numerical solution algorithm for related problem is given in the paper [71], the general results for the algorithm is described in [72]).

The main advantages of the chosen method are effectiveness for high dimensions, parallelization or even distributed evaluation, availability for categorical variables and discontinuous functions, ease of implementation. The limitations are the weak convergence guarantees, being surrogate modeling, disregard of properties of objective function, and lack of uncertainty estimations.

The convergence results for Novosibirsk region are presented in the Figure A2.

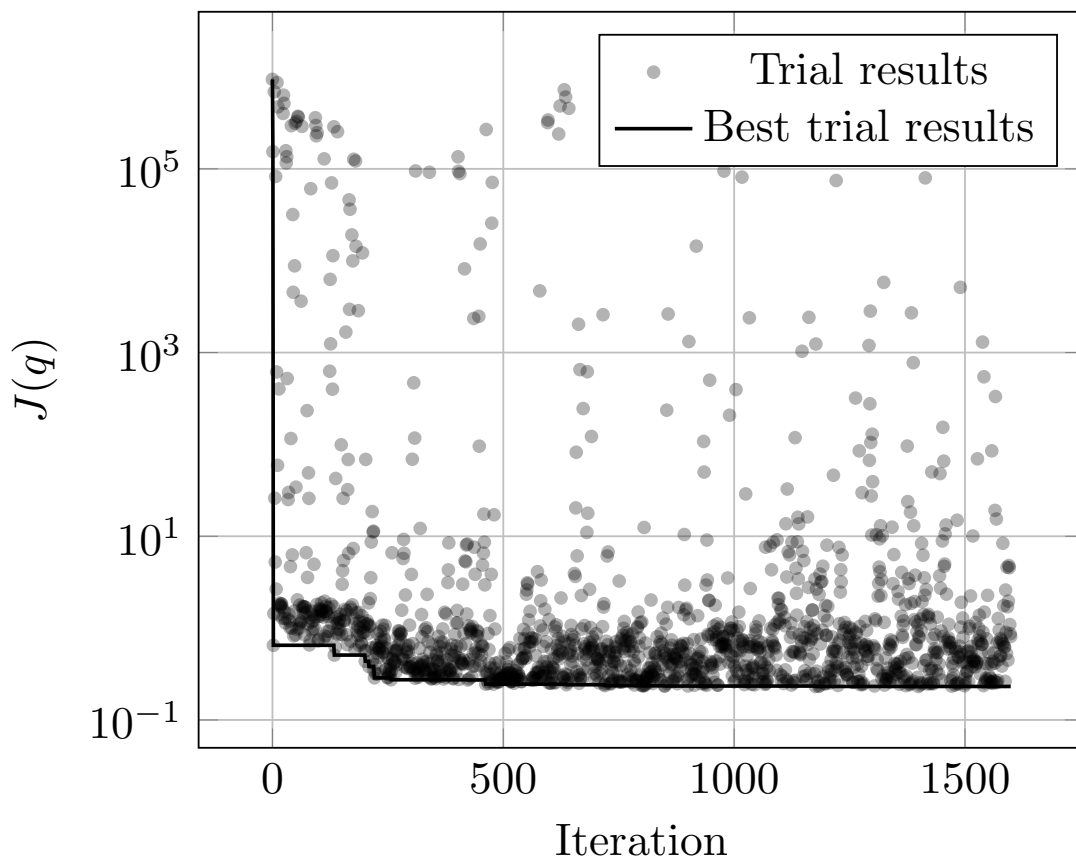


Figure A2. Tree-based Parzen estimator convergence results for Novosibirsk region.

Appendix A.5. The Pareto Front of the Inverse Problem Solution

The Pareto front for TPE optimization is shown in Figure A3.

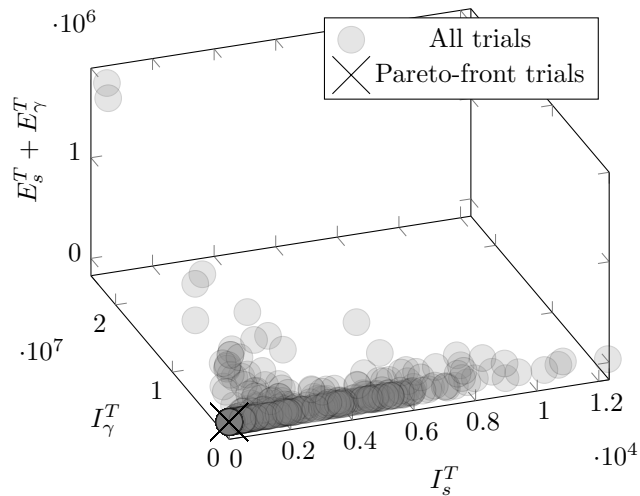


Figure A3. Pareto front for Altai republic optimization with TPE algorithm.

Appendix A.6. 2019 Year Data Spike

The effect of 2019 spike in data is shown in Figure A4. In comparison to the results in Figure 5, the difference in results is not big, however, the 2019 data anomaly is not describable with differential model, thus it is removed from the dataset.

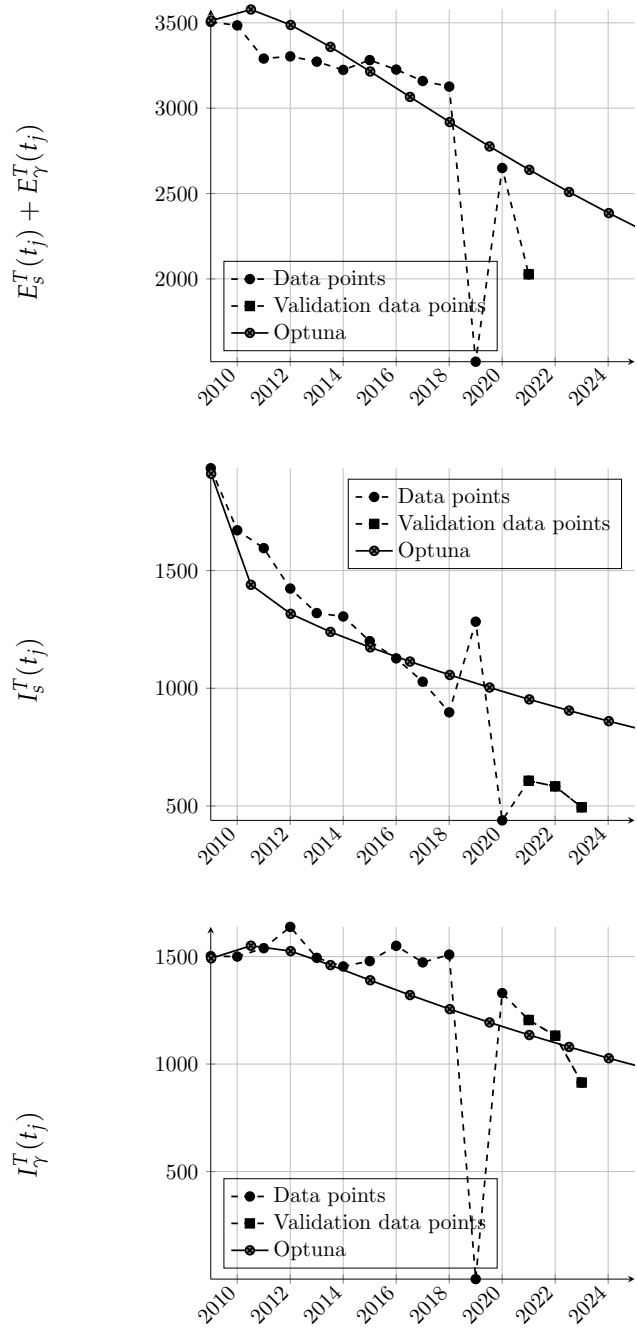


Figure A4. Restored forecast of infected individuals in the Novosibirks region, compared with actual data (dash-dot lines) with 2019 year data.

References

1. Global Tuberculosis Report, 2024. (accessed 2024-11-25).
2. Goletti, D.; Meintjes, G.; Andrade, B.; Zumla, A.; Lee, S.s. Insights from the 2024 WHO Global Tuberculosis Report – More Comprehensive Action, Innovation, and Investments required for achieving WHO End TB goals. *International Journal of Infectious Diseases* **2024**, *150*, 107325. <https://doi.org/10.1016/j.ijid.2024.107325>.
3. Hamna Mariyam, K.; Sayooj Aby, J.; Anuwat, J.; Karuna, M. A comprehensive study on tuberculosis prediction models: Integrating machine learning into epidemiological analysis. *Journal of Theoretical Biology* **2025**, *597*, 111988. <https://doi.org/10.1016/j.jtbi.2024.111988>.
4. Krivorotko, O.; Kabanikhin, S. Artificial intelligence for COVID-19 spread modeling. *Journal of Inverse and Ill-posed Problems* **2024**, *32*, 297 – 332. <https://doi.org/10.1515/jiip-2024-0013>.

5. Avilov, K.; Romanukha, A. Mathematical models of spread and control of tuberculosis (review) (in Russian). *Mathematical biology and bioinformatics* **2007**, 2, 188–318. <https://doi.org/10.17537/2007.2.188>.
6. Melnichenko, O.; Romanukha, A. Model of the tuberculosis epidemiology. Data analysis and parameters estimation (in Russian). *Mathematical modelling* **2008**, 20, 107–128. <https://doi.org/10.1134/S2070048209040024>.
7. Kabanikhin, S.; Krivorotko, O.; Neverov, A.; Kaminskiy, G.; Semenova, O. Identification of the Mathematical Model of Tuberculosis and HIV Co-Infection Dynamics. *Mathematics* **2024**, 12. <https://doi.org/10.3390/math12233636>.
8. Vlad, A.I.; Romanyukha, A.A.; Sannikova, T.E. Parameter Tuning of Agent-Based Models: Metaheuristic Algorithms. *Mathematics* **2024**, 12. <https://doi.org/10.3390/math12142208>.
9. Romanyukha, A.A.; Karkach, A.S.; Borisov, S.; Belilovsky, E.M.; Sannikova, T.E.; Krivorotko, O. Small-scale stable clusters of elevated tuberculosis incidence in Moscow, 2000–2015. Discovery and spatio-temporal analysis. *International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases* **2019**. <https://doi.org/10.1016/j.ijid.2019.11.015>.
10. Krivorotko, O.I.; Zyatkov, N.Y.; Kabanikhin, S.I. Epidemics modelling: neural net based on data and SIR-model (in Russian). *Computational Mathematics and Mathematical Physics* **2023**, 63, 1733–1746. <https://doi.org/10.31857/S0044466923100095>.
11. Krivorotko, O.; Zyatkov, N. The Forecasting of the Spread of Infectious Diseases Based on Conditional Generative Adversarial Networks. *Mathematics* **2024**, 12. <https://doi.org/10.3390/math12193044>.
12. Meshkov, I.A.; Petrenko, T.; Keiser, O.; Estill, J.; Revyakina, O.V.; Felker, I.; Raviglione, M.C.; Krasnov, V.A.; Schwartz, Y. Variations in tuberculosis prevalence, Russian Federation: a multivariate approach. *Bulletin of the World Health Organization* **2019**, 97, 737 – 745A. <https://doi.org/10.2471/BLT.19.229997>.
13. Drain, P.K.; Bajema, K.L.; Dowdy, D.W.; Dheda, K.; Naidoo, K.; Schumacher, S.G.; Ma, S.; Meemeier, E.W.; Lewinsohn, D.M.; Sherman, D.R. Incipient and Subclinical Tuberculosis: a Clinical Review of Early Stages and Progression of Infection. *Clinical Microbiology Reviews* **2018**, 31. <https://doi.org/10.1128/CMR.00021-18>.
14. Janssen, S.; Murphy, M.; Upton, C.; Allwood, B.; Diacon, A.H. Tuberculosis: An Update for the Clinician. *Respirology* **2025**, 30, 196–205, [<https://onlinelibrary.wiley.com/doi/pdf/10.1111/resp.14887>]. <https://doi.org/10.1111/resp.14887>.
15. D.W. Dowdy, S. Basu, J.A. Is passive diagnosis enough? The impact of subclinical disease on diagnostic strategies for tuberculosis. *Am. J. Respir. Crit. Care Med* **2013**, 187, 543–551. <https://doi.org/10.1164/rccm.201207-1217OC>.
16. Calligaro, G.L.; Zijenah, L.S.; Peter, J.; Theron, G.; Buser, V.; McNerney, R.; Bara, W.; Bandason, T.; Govender, U.; Tomasicchio, M.; et al. Effect of new tuberculosis diagnostic technologies on community-based intensified case finding: a multicentre randomised controlled trial. *The Lancet. Infectious diseases* **2017**, 17 4, 441–450. [https://doi.org/10.1016/S1473-3099\(16\)30384-X](https://doi.org/10.1016/S1473-3099(16)30384-X).
17. Sabiiti, W.; Mtafya, B.; de Lima, D.A.; Dombay, E.; Baron, V.O.; Azam, K.; Oravcová, K.; Sloan, D.J.; Gillespie, S.H. A Tuberculosis Molecular Bacterial Load Assay (TB-MBLA). *Journal of visualized experiments : JoVE* **2020**, 158. <https://doi.org/10.3791/60460>.
18. Gillespie, S.H.; Sabiiti, W. Tuberculosis molecular bacterial load assay in the management of tuberculosis. *Current Opinion in Infectious Diseases* **2024**, 38, 176 – 181. <https://doi.org/10.1097/QCO.0000000000001084>.
19. Wang, J.; xia Zhang, X.; Huo, F.; Qin, L.; Liu, R.; Shang, Y.; Yao, C.; Ma, L.; Pang, Y. Analysis of Xpert MTB/RIF results in retested patients with very low initial bacterial loads: A retrospective study in China. *Journal of infection and public health* **2023**, 16 6, 911–916. <https://doi.org/10.1016/j.jiph.2023.04.004>.
20. Mbelele, P.M.; Sabiiti, W.; Heysell, S.K.; Sauli, E.; Mpolya, E.A.; Mfinanga, S.; Gillespie, S.H.; Addo, K.K.; Kibiki, G.S.; Sloan, D.J.; et al. Use of a molecular bacterial load assay to distinguish between active TB and post-TB lung disease. *The International Journal of Tuberculosis and Lung Disease* **2022**, 26, 276 – 278. <https://doi.org/10.5588/ijtld.21.0459>.
21. Din, A. Bifurcation analysis of a delayed stochastic HBV epidemic model: Cell-to-cell transmission. *Chaos, Solitons and Fractals* **2024**, 181, 114714. <https://doi.org/10.1016/j.chaos.2024.114714>.
22. DIN, A.; Li, Y. Optimizing HIV/AIDS dynamics: stochastic control strategies with education and treatment. *European Physical Journal Plus* **2024**, 139, 812. <https://doi.org/10.1140/epjp/s13360-024-05605-1>.
23. Kim, P.S.; Swaminathan, S. Ending TB: the world's oldest pandemic. *Journal of the International AIDS Society* **2021**, 24, e25698. <https://doi.org/10.1002/jia2.25698>.
24. Wallace, D.; Wallace, R. Problems with the WHO TB model. *Mathematical Biosciences* **2019**, 313, 71–80. <https://doi.org/10.1016/j.mbs.2019.05.002>.

25. Xu, C.; Cheng, K.; Wang, Y.; Liu, M.; Wang, X.; Yang, Z.; Guo, S. Analysis of the current status of TB transmission in China based on an age heterogeneity model. *Mathematical Biosciences and Engineering* **2023**, *20*, 19232–19253. <https://doi.org/10.3934/mbe.2023850>.
26. Lin, Y.J.; Liao, C.M. Seasonal dynamics of tuberculosis epidemics and implications for multidrug-resistant infection risk assessment. *Epidemiology and Infection* **2014**, *142*, 358–370. <https://doi.org/10.1017/S0950268813001040>.
27. H.X.Wang.; L.Q.Jiang.; G.H.Wang. Stability of a tuberculosis model with a time delay in transmission. *Journal of North University of China* **2014**, *35*, 238–242. <https://doi.org/10.3969/j.issn.1673-3193.2014.03.003>.
28. Achkar, J.M.; Jenny-Avital, E.R. Incipient and Subclinical Tuberculosis: Defining Early Disease States in the Context of Host Immune Response. *Journal of Infectious Diseases* **2011**, *204*, S1179–S1186. <https://doi.org/10.1093/infdis/jir451>.
29. Migliori, G.B.; Ong, C.W.; Petrone, L.; D'Ambrosio, L.; Centis, R.; Goletti, D. The definition of tuberculosis infection based on the spectrum of tuberculosis disease. *Breathe* **2021**, *17*, 210079. <https://doi.org/10.1183/20734735.0079-2021>.
30. Zaidi, S.M.; Coussens, A.K.; Seddon, J.A.; Kredo, T.; Warner, D.; Houben, R.M.; Esmail, H. Beyond latent and active tuberculosis: a scoping review of conceptual frameworks. *eClinicalMedicine* **2023**, *66*, 102332. <https://doi.org/10.1016/j.eclinm.2023.102332>.
31. Herrera, M.; Taguiam, E.; Laupland, K.B.; Rueda, Z.V.; Keynan, Y. Public health implications of the evolving understanding of tuberculosis natural history. *Journal of the Association of Medical Microbiology and Infectious Disease Canada* **2023**, *8*, 241–244. <https://doi.org/10.3138/jammi-2023-02-24>.
32. Emery, J.C.; Dodd, P.J.; Banu, S.; Frascella, B.; Garden, F.L.; Horton, K.C.; Hossain, S.; Law, I.; Van Leth, F.; Marks, G.B.; et al. Estimating the contribution of subclinical tuberculosis disease to transmission: An individual patient data analysis from prevalence surveys. *eLife* **2023**, *12*, e82469. <https://doi.org/10.7554/eLife.82469>.
33. Escalante, P.; Vadiyala, M.R.; Pathakumari, B.; Marty, P.K.; Van Keulen, V.P.; Hilgart, H.R.; Meserve, K.; Theel, E.S.; Peikert, T.; Bailey, R.C.; et al. New diagnostics for the spectrum of asymptomatic TB: from infection to subclinical disease. *The International Journal of Tuberculosis and Lung Disease* **2023**, *27*, 499–505. <https://doi.org/10.5588/ijtld.23.0032>.
34. Mahmoudi, S.; García, M.J.; Drain, P.K. Current approaches for diagnosis of subclinical pulmonary tuberculosis, clinical implications and future perspectives: a scoping review. *Expert Review of Clinical Immunology* **2024**, *20*, 715–726. <https://doi.org/10.1080/1744666X.2024.2326032>.
35. Cho, S.; Shin, S.; Kim, Y.; Song, W.; Hong, S.; Jeong, S.; Kang, M.; Lee, K. A novel approach for tuberculosis diagnosis using exosomal DNA and droplet digital PCR. *Clinical Microbiology and Infection* **2020**, *26*, 942.e1–942.e5. <https://doi.org/10.1016/j.cmi.2019.11.012>.
36. Gupta, R.K.; Turner, C.T.; Venturini, C.; Esmail, H.; Rangaka, M.X.; Copas, A.; Lipman, M.; Abubakar, I.; Noursadeghi, M. Concise whole blood transcriptional signatures for incipient tuberculosis: a systematic review and patient-level pooled meta-analysis. *The Lancet Respiratory Medicine* **2020**, *8*, 395–406. [https://doi.org/10.1016/S2213-2600\(19\)30282-6](https://doi.org/10.1016/S2213-2600(19)30282-6).
37. Rotundo, S.; Tassone, M.T.; Serapide, F.; Russo, A.; Trecarichi, E.M. Incipient tuberculosis: a comprehensive overview. *Infection* **2024**, *52*, 1215–1222. <https://doi.org/10.1007/s15010-024-02239-4>.
38. Sivakumaran, D.; Jenum, S.; Srivastava, A.; Steen, V.M.; Vaz, M.; Doherty, T.M.; Ritz, C.; Grewal, H.M.S. Host blood-based biosignatures for subclinical TB and incipient TB: A prospective study of adult TB household contacts in Southern India. *Frontiers in Immunology* **2023**, *13*, 1051963. <https://doi.org/10.3389/fimmu.2022.1051963>.
39. Perelman, M.; Marchuk, G.; Borisov, S.; Kazennykh, B.; Avilov, K.; Karkach, A.; Romanyukha, A. Tuberculosis epidemiology in Russia: the mathematical model and data analysis. *Russian Journal of Numerical Analysis and Mathematical Modelling - RUSS J NUMER ANAL MATH MODEL* **2004**, *19*, 305–314. <https://doi.org/10.1515/1569398041974905>.
40. Yu, Y.; Shi, Y.; Yao, W. Dynamic model of tuberculosis considering multi-drug resistance and their applications. *Infectious Disease Modelling* **2018**, *3*, 362–372. <https://doi.org/10.1016/j.idm.2018.11.001>.
41. Avilov, K.; Romanukha, A.; Belilovsky, E.; Borisov, S. Mathematical modelling of the progression of active tuberculosis: Insights from fluorography data. *Infectious Disease Modelling* **2022**, *7*, 374–386. <https://doi.org/10.1016/j.idm.2022.06.007>.
42. Unified Interdepartmental Information and Statistical System. <https://www.fedstat.ru/>. Accessed: 2026-19-13.

43. Stavitskaya, N.; Nemkova, E.; Tashkova, G. *Key indicators of anti-tuberculosis activities in the Siberian and Far Eastern Federal Districts (statistical materials) (in russian)*; Novosibirsk TB Research Institute: Novosibirsk, Russia, 2025; p. 156. <https://doi.org/10.3390/math12233636>.
44. Nguyen, H.V.; Tiemersma, E.; Nguyen, N.V.; Nguyen, H.B.; Cobelens, F. Disease Transmission by Patients With Subclinical Tuberculosis. *Clinical Infectious Diseases* **2023**, *76*, 2000–2006. <https://doi.org/10.1093/cid/ciad027>.
45. Stavitskaya N.V, Nemkova E.K., T.G. Key indicators of anti-tuberculosis activities in the Siberian and Far Eastern Federal Districts (statistical materials) (in russian), 2025. <https://doi.org/https://elibrary.ru/item.asp?id=82569907>.
46. Tyulkova, T.Y.; Mezentseva, A.V. Latent Tuberculosis Infection and Residual Post-Tuberculous Changes in Children. *Current pediatrics* **2017**, *16*, 452–456. <https://doi.org/10.15690/vsp.v16i6.1817>.
47. Tornack, J.; Reece, S.T.; Bauer, W.M.; Vogelzang, A.; Bandermann, S.; Zedler, U.; Stingl, G.; Kaufmann, S.H.E.; Melchers, F. Human and Mouse Hematopoietic Stem Cells Are a Depot for Dormant Mycobacterium tuberculosis. *PLOS ONE* **2017**, *12*, e0169119. <https://doi.org/10.1371/journal.pone.0169119>.
48. Belogorodtsev, S.N.; Lykov, A.P.; Nemkova, E.K.; Schwartz, Y.S. Interaction between mesenchymal stromal cells and tuberculous mycobacteria in vitro. *Tuberculosis and Lung Diseases* **2023**, *101*, 57–63. <https://doi.org/10.58838/2075-1230-2023-101-1S-57-63>.
49. Khan, A.; Jagannath, C. Interactions of Mycobacterium tuberculosis with Human Mesenchymal Stem Cells. In *Tuberculosis Host-Pathogen Interactions*; Cirillo, J.D.; Kong, Y., Eds.; Springer International Publishing: Cham, 2019; pp. 95–111. https://doi.org/10.1007/978-3-030-25381-3_5.
50. Dale, K.D.; Schwalb, A.; Coussens, A.K.; Gibney, K.B.; Abboud, A.J.; Watts, K.; Denholm, J.T. Overlooked, dismissed, and downplayed: reversion of *Mycobacterium tuberculosis* immunoreactivity. *European Respiratory Review* **2024**, *33*, 240007. <https://doi.org/10.1183/16000617.0007-2024>.
51. Shang, Y. Lie algebraic discussion for affinity based information diffusion in social networks. *Open Physics* **2017**, *15*, 705–711. <https://doi.org/10.1515/phys-2017-0083>.
52. Demography statistics of Russian Federation by Rosstat, 2025, <https://rosstat.gov.ru/folder/12781>. (accessed 2025-12-01).
53. van den Driessche, P.; Watmough, J., Further Notes on the Basic Reproduction Number. In *Mathematical Epidemiology*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2008; pp. 159–178. https://doi.org/10.1007/978-3-540-78911-6_6.
54. Kabanikhin, S.I.; Voronov, D.A.; Grodz, A.A.; Krivorotko, O.I. Identifiability of mathematical models in medical biology. *Russian Journal of Genetics: Applied Research* **2016**, *6*, 838–844. <https://doi.org/10.1134/S2079059716070054>.
55. Krivorotko, O.; Kabanikhin, S.; Petrakova, V. Identifiability of mathematical models of epidemiology: tuberculosis, HIV, COVID-19. *Mathematical biology and bioinformatics* **2023**, pp. 177–214. <https://doi.org/10.17537/2023.18.177>.
56. Sobol, I. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *MATH COMPUT SIMULAT* **2001**, *55*, 271–280. [https://doi.org/10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6).
57. Order On approval of methodological recommendations for improving the diagnosis and treatment of tuberculosis of the respiratory organs, 2014. (dated 29.12.2014 No. 951).
58. Federal Law No. 314-FZ dated August 3, 2018, 2018.
59. Order of the Ministry of Health of the Russian Federation dated 13.03.2019 N 127n "On Approval of the Procedure for Dispensary Monitoring of Tuberculosis Patients, Persons in or Who Have Been in Contact with a Source of Tuberculosis, as Well as Persons with Suspected Tuberculosis and Those Who Have Been Treated for Tuberculosis, and on the Repeal of Items 16 - 17 of the Procedure for Providing Medical Care to Tuberculosis Patients, Approved by Order of the Ministry of Health of the Russian Federation dated 15.11.2012 N 932n.
60. Order of the Russian Ministry of Health dated 05.04.2019 N 199 "On Approval of the Departmental Target Program "Prevention and Control of Socially Significant Infectious Diseases".
61. Trajman, A.; Felker, I.; Alves, L.C.; Coutinho, I.; Osman, M.; Meehan, S.A.; Singh, U.B.; Schwartz, Y. The COVID-19 and TB syndemic: the way forward. *The International Journal of Tuberculosis and Lung Disease* **2022**, *26*, 710–719. <https://doi.org/10.5588/ijtld.22.0006>.
62. Avilov, K.; Romanukha, A.; Belilovsky, E.; Borisov, S. Comparison of simulation schemes for the natural history of respiratory tuberculosis. *Mathematical biology and bioinformatics* **2019**, *14*, 570–587. <https://doi.org/10.17537/2019.14.570>.

63. Nguyen, H.; Tiemersma, E.; Nguyen, N.; Nguyen, H.; Cobelens, F. Disease Transmission by Patients With Subclinical Tuberculosis. *Clinical infectious diseases* **2023**, *76*, 2000–2006. <https://doi.org/10.1093/cid/ciad027>.
64. Guo, Z.K.; Huo, H.F.; Xiang, H.; Ren, Q.Y. Global dynamics of a tuberculosis model with age-dependent latency and time delays in treatment. *Journal of Mathematical Biology* **2023**, *87*, 66. <https://doi.org/10.1007/s00285-023-01999-1>.
65. Ochieng, F.O. SEIRS model for TB transmission dynamics incorporating the environment and optimal control. *BMC Infectious Diseases* **2025**, *25*, 490. <https://doi.org/10.1186/s12879-025-10710-2>.
66. Yu, Y.; Shi, Y.; Yao, W. Dynamic model of tuberculosis considering multi-drug resistance and their applications. *Infectious Disease Modelling* **2018**, *3*, 362–372. <https://doi.org/10.1016/j.idm.2018.11.001>.
67. Villalva-Serra, K.; Barreto-Duarte, B.; Rodrigues, M.M.; Queiroz, A.T.; Martinez, L.; Croda, J.; Rolla, V.C.; Kritski, A.L.; Cordeiro-Santos, M.; Sterling, T.R.; et al. Impact of strategic public health interventions to reduce tuberculosis incidence in Brazil: a Bayesian structural time-series scenario analysis. *The Lancet Regional Health - Americas* **2025**, *41*, 100963. <https://doi.org/10.1016/j.lana.2024.100963>.
68. Villalva-Serra, K.; Barreto-Duarte, B.; Rodrigues, M.M.; Queiroz, A.T.; Martinez, L.; Croda, J.; Rolla, V.C.; Kritski, A.L.; Cordeiro-Santos, M.; Sterling, T.R.; et al. Corrigendum to “Impact of strategic public health interventions to reduce tuberculosis incidence in Brazil: a Bayesian structural time-series scenario analysis” - The Lancet Regional Health – Americas 2025; Volume 41: 100963; DOI: 10.1016/j.lana.2024.100963. *The Lancet Regional Health - Americas* **2026**, *54*, 101381. <https://doi.org/10.1016/j.lana.2026.101381>.
69. Saber, S.; Alahmari, A. Mathematical Insights into Zoonotic Disease Spread: Application of the Milstein Method. *European Journal of Pure and Applied Mathematics* **2025**, *18*, 5881. <https://doi.org/10.29020/nybg.ejpam.v18i2.5881>.
70. Saber, S.; Solouma, E.; Althubiani, M.; Messaoudi, M. Statistical Insights into Zoonotic Disease Dynamics: Simulation and Control Strategy Evaluation. *Symmetry* **2025**, *17*, 733. <https://doi.org/10.3390/sym17050733>.
71. Kabanikhin, S.I.; Krivorotko, O.; Neverov, A.; Kaminskiy, G.; Semenova, O. Identification of the Mathematical Model of Tuberculosis and HIV Co-Infection Dynamics. *Mathematics* **2024**. <https://doi.org/10.3390/math12233636>.
72. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A Next-generation Hyperparameter Optimization Framework, 2019, [arXiv:cs.LG/1907.10902]. <https://doi.org/10.48550/arXiv.1907.10902>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.